TEHNIČKI GLASNIK

TECHNICAL JOURNAL

TEHNIČKI GLASNIK / TECHNICAL JOURNAL – GODIŠTE / VOLUME 20 – BROJ / ISSUE 2

LIPANJ 2026 / JUNE 2026 – STRANICA / PAGES 173-342

/

SVEUČILIŠTE SJEVER / UNIVERSITY NORTH – CROATIA – EUROPE

ISSN 1846-6168 (PRINT) / ISSN 1848-5588 (ONLINE)

ISSN 1846-6168 (Print) ISSN 1848-5588 (Online)

TEHNIČKI GLASNIK - TECHNICAL JOURNAL

Scientific-professional journal of University North

Volume 20 Issue 2
Varaždin, June 2026 Pages 173-342

Editorial Office:
Sveučilište Sjever / University North – Tehnički glasnik / Technical journal
Sveučilišni centar Varaždin / University Center Varaždin
Jurja Križanića 31b, 42000 Varaždin, Croatia
Tel. ++385 42 493 338, Fax.++385 42 493 336
E-mail: tehnickiglasnik@unin.hr
https://tehnickiglasnik.unin.hr
https://www.unin.hr/djelatnost/izdavastvo/tehnicki-glasnik/
https://hrcak.srce.hr/tehnickiglasnik

Founder and Publisher:
Sveučilište Sjever / University North

Council of Journal:
Damir VUSIĆ, Chairman; Marin MILKOVIĆ, Member; Robert GEČEK, Member; Mario TOMIŠA, Member;
Marijana NEUBERG, Member; Danko MARKOVINOVIĆ, Member; Petar MIŠEVIĆ, Member

Editorial Board:
Domestic Members:
Chairman Marin MILKOVIĆ (1), Milan KLJAJIN (1), Damir VUSIĆ (1), Krešimir BUNTAK (1), Anica HUNJET (1), Živko KONDIĆ (1), Goran KOZINA (1), Ljudevit KRPAN (1), Krunoslav HAJDEK (1), Marko STOJIĆ (1), Božo SOLDO (1), Mario TOMIŠA (1), Vlado TROPŠA (1), Vinko VIŠNJIĆ (1), Sanja ŠOLIĆ (1), Dean VALDEC (1), Predrag PUTNIK (1), Petar MIŠEVIĆ (1), Duško PAVLETIĆ (5), Branimir PAVKOVIĆ (5), Mile MATIJEVIĆ (3), Damir MODRIĆ (3), Nikola MRVAC (3), Klaudio PAP (3), Ivana ŽILJAK STANIMIROVIĆ (3), Krešimir GRILEC (6), Biserka RUNJE (6), Sara HAVRLIŠAN (2), Dražan KOZAK (2), Roberto LUJIĆ (2), Leon MAGLIĆ (2), Ivan SAMARDŽIĆ (2), Antun STOIĆ (2), Katica ŠIMUNOVIĆ (2), Goran ŠIMUNOVIĆ (2), Ladislav LAZIĆ (7), Ante ČIKIĆ (1), Darko DUKIĆ (9), Gordana DUKIĆ (10), Srđan MEDIĆ (11), Sanja KALAMBURA (12), Marko DUNĐER (13), Zlata DOLAČEK-ALDUK (4), Dina STOBER (4)
International Members:
Boris TOVORNIK (14), Milan KUHTA (15), Nenad INJAC (16), Marin PETROVIĆ (18), Salim IBRAHIMEFENDIĆ (19), Zoran LOVREKOVIĆ (20), Igor BUDAK (21), Darko BAJIĆ (22), Tomáš HANÁK (23), Evgenij KLIMENKO (24), Oleg POPOV (24), Ivo ČOLAK (25), Katarina MONKOVÁ (26), Berenika HAUSNEROVÁ (8), Nenad GUBELJAK (27), Stefanija KLARIC (28), Bertrand MARESCHAL (29), Sachin R. SAKHARE (30), Suresh LIMKAR (31), Mandeep KAUR (32), Aleksandar SEDMAK (33), Han-Chieh CHAO (34), Sergej HLOCH (26), Grzegorz M. KRÓLCZYK (35), Djordje VUKELIC (21), Stanisław LEGUTKO (17), Valentin POPOV (36), Dragan MARINKOVIC (36), Hamid M. SEDIGHI (37), Cristiano FRAGASSA (38), Dragan PAMUČAR (39), Imre FELDE (40), Levente KOVACS (40)

Editor-in-Chief:
Milan KLJAJIN

Technical Editor:
Milan KLJAJIN

Graphics Editor:
Snježana IVANČIĆ VALENKO

IT support:
Antonija MANDIĆ

Print:
Centar za digitalno nakladništvo, Sveučilište Sjever

All manuscripts published in journal have been reviewed.
Manuscripts are not returned.

The journal is free of charge and four issues per year are published
(In March, June, September and December)
Circulation: 100 copies

Journal is indexed and abstracted in:
Web of Science Core Collection (Emerging Sources Citation Index - ESCI), Scopus, EBSCOhost Academic Search Complete, EBSCOhost – One Belt, One Road Reference Source Product, ERIH PLUS, CITEFACTOR – Academic Scientific Journals, DOAJ – Directory of Open Access Journals, Hrčak – Portal znanstvenih časopisa RH

Registration of journal:
The journal "Tehnički glasnik" is listed in the HGK Register on the issuance and distribution of printed editions on the 18th October 2007 under number 825.

Preparation ended:	Published (online):	Published (print):
April 13, 2026	April 17, 2026	June 15, 2026

Legend:
(1) University North, (2) University of Slavonski Brod, (3) Faculty of Graphic Arts Zagreb, (4) Faculty of Civil Engineering Osijek, (5) Faculty of Engineering Rijeka, (6) Faculty of Mechanical Engineering and Naval Architecture Zagreb, (7) Faculty of Metallurgy Sisak, (8) Tomas Bata University in Zlín, (9) Department of Physics of the University of Josip Juraj Strossmayer in Osijek, (10) Faculty of Humanities and Social Sciences Osijek, (11) Karlovac University of Applied Sciences, (12) University of Applied Sciences Velika Gorica, (13) Department of Polytechnics - Faculty of Humanities and Social Sciences Rijeka, (14) Faculty of Electrical Engineering and Computer Science - University of Maribor, (15) Faculty of Civil Engineering - University of Maribor, (16) University College of Teacher Education of Christian Churches Vienna/Krems, (17) Faculty of Mechanical Engineering - Poznan University of Technology (Poland), (18) Mechanical Engineering Faculty Sarajevo, (19) University of Travnik - Faculty of Technical Studies, (20) Higher Education Technical School of Professional Studies in Novi Sad, (21) University of Novi Sad - Faculty of Technical Sciences, (22) Faculty of Mechanical Engineering - University of Montenegro, (23) Brno University of Technology, (24) Odessa State Academy of Civil Engineering and Architecture, (25) Faculty of Civil Engineering - University of Mostar, (26) Faculty of Manufacturing Technologies with the seat in Prešov - Technical University in Košice, (27) Faculty of Mechanical Engineering - University of Maribor, (28) College of Engineering, IT & Environment - Charles Darwin University, (29) Universite Libre de Bruxelles, (30) Vishwakarma Institute of Information Technology (Pune, India), (31) AISSMS Institute of Information Technology (Pune, India), (32) Permtech Research Solutions (India), (33) University of Belgrade, (34) National Dong Hwa University - Taiwan, (35) Faculty of Mechanical Engineering - Opole University of Technology (Poland), (36) TU Berlin - Germany, (37) Shahid Chamran University of Ahvaz - Iran, (38) University of Bologna - Italy, (39) University of Defence in Belgrade - Military Academy – Serbia, (40) Obuda University Budapest - Hungary

© 2026 **Tehnički glasnik / Technical Journal.** All rights reserved

CONTENT **I**

Mirche Stojkov*, Zoran Pandilov
Neuro Fuzzy – Regression Based Concept for Dosing Determination **173**

Badeea Al Sukhni, Bhabendu Kumar Mohanta*, Debnath Bhattacharyya, Tai-hoon Kim*
Investigate Unsolicited Traffic on IoT Devices Using Machine Learning **179**

Reza Askarizad*, Roomina Soleymani Ardejani, Hossein Safari
Empirical Assessment of Audience Attraction in Shopping Centres Based on the Notion of Social Sustainability **188**

Somayye Nasiri, Shahram Shahabi, Aynaz Shafiesabet, Maral Talebbeidokhti*, Elmira Alvandi Behineh
Cybersecurity in Action: Unraveling the Effects of Individual, Social, and Organizational Determinants **195**

Petar Pejić*, Jelena Pejić
Accelerating Game Level Design with Machine Learning: A Unity Module for Procedural Kitchen Generation **205**

Samuel Lakkala*, Joonas Suutari, Pekka Tervonen, Harri Haapasalo
Customer Relationship Management Processes and Practices in a Higher Education Institution: A Case Study **210**

Kazem Mehrabi, Mohammad Farhadi-Kangarlu*, Behrouz Tousi
An Adaptive Load-Shedding Approach Applying Wide-Area Measurement Systems (WAMS) **218**

Asiye Omidian, Athena Omidian*, Shahriar Mohajer, Elnaz Zarei, Azita Ramezani
Sustainable Supply Chain Management, Information Technology, and Firm Performance: The Mediating Power of Agility and Supply Chain Integration **226**

Yeong-Chae Yun, Seok-Heon Yun*
Hierarchical Ensemble Learning for Automatic Classification of Construction Work Codes **233**

Sang Hyun Yoo, Hyun Jung Kim*
Adaptive GANs with Domain-Specific Losses for Data Imbalance **240**

Andreja Rudančić*, Petar Mišević, Sandra Šokčević
Impacts of Communication on Interpersonal Relationships in Business Management **249**

Milner Liendo*, Manuel Hilario, Alonso Tenorio, Jaime Cancho
Technological Innovation and Circular Economy: Smart Containers and Real-Time Monitoring for Efficient Waste Management **258**

Sara Havrlišan*, Martina Tokić, Štefanija Klarić, Branko Štrbac
Influence of the Process Parameters on the Mechanical Properties of the FDM 3D Printed Samples **269**

Kailas Deore*, Nilesh Salunke, Sudhakar Pawar, Yogeshkumar Pathak, Juber Khatik
Experimental Performance Analysis of PCM and Fin-Integrated Solar Stills under Varying Water Depths **277**

Davronbek Malikov, Jaeho Kim*
Machine Learning-Based xG forecasting in the Top 5 European Soccer Leagues: A Comparative Analysis **286**

Rehab I. Ajel*, Tariq A. Hassan
Speech Signal Enhancement Using Linear Predictive Coefficients Adaptive Filter **297**

Juuso Kivijakola*, Silja Heikkinen, Pekka Tervonen, Harri Haapasalo
Utilising Data Centres' Side Streams – A Waste Heat Case Study **306**

Pejo Konjatić, Josip Kačmarčik*, Marko Katinić, Lamija Mešeljević
Tensile Properties Regression Analysis of FDM 3D Printed Composites with ONYX Matrix and Different Fibres **314**

Mateo Vujčić, Tomislav Horvat*, Dejan Barešić, Dražen Crčić
Supervised Machine Learning Methods in Predicting English Premier League Game Outcomes **321**

Marcel Piskač
A Comprehensive Review of Artificial Intelligence Applications in Financial Market Predictions **328**

Igor Majnarić, Marko Morić*, Marko Kadija, Krunoslav Hajdek
Realization Lines with Inkjet Head KM1024iLHE-30 and Additional Applying Metalized Foil **334**

INSTRUCTIONS FOR AUTHORS **III**


WIT
CONFERENCES
Call for Papers
BEM/MRM
48
16-18 October 2026 | Dalian, China
Organised by
Dalian University of Technology, CHINA
Conference Affiliated with
Wessex Institute of Technology, UK
Sponsored by
State Key Laboratory of Structural Analysis, Optimization and CAE Software for Industrial Equipment, CHINA
https://BEM-MRM48.dlut.edu.cn

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20230713121334

Original scientific paper
Received: 2023-07-13, Accepted: 2025-09-08

Neuro Fuzzy – Regression Based Concept for Dosing Determination

Mirche Stojkov*, Zoran Pandilov

Abstract: This research aims to optimize dosing through the application of intelligent tools-methods. Here, the case study focuses on the flocculation phase of drinking water processing in a real system. Adaptive Neuro Fuzzy Inference System (ANFIS), regression, and "know-how" of the operators represented through linguistic expression as experiential knowledge are integrated to create a dosing calculation program. The goal is to establish a functional relationship between the process variables.

Keywords: Adaptive Neuro Fuzzy Inference System (ANFIS); dosing; experiential knowledge; linguistic expression; regression

1 INTRODUCTION

When treating drinking water, it is important to meet two key requirements: to ensure that the water is safe for consumption and to minimize the costs of the treatment process. Overdose is increasing costs, while insufficient dosing disrupts water treatment process. By using the correct amount of coagulant-flocculent, two important goals are achieved. The first objective is to improve the efficiency of flocculation, which is an important step for water treatment. The second objective is to reduce the costs associated with the process. Typically, operators rely on their experience to determine the appropriate amount of dosing in processing facilities.

The JAR test it simulates the coagulation-flocculation and sedimentation process in a water treatment plant and helps operators determine if they are using the right amount of treatment chemicals. However, in reality, the determination is delayed because it relies on the already changed properties of the raw water. Operators use their experience to verbally group certain values of the variables noted in the worksheets. In past in most cases, intelligent tools-methods were used for the purpose of optimization and modelling in water processing systems. This research integrates the use of these methods. In addition, these methods have application in other various fields. In researches [1-9] are shown the use of methods: ANN, Fuzzy Logic and Regression like valuable tools for optimization, prediction and computing. In a similar studies like our research, input data from surface water and operational data from dosage, are used to create models for the dosing prediction and control also using: ANN, Fuzzy Logic and Regression Analysis [10-16]. The primary goal of this study is to create a model (program) that could serve as a control unit. In this study, flow and turbidity of raw water and dosage of flocculent solution are used as input-output variables. To establish the model, 240 pairs or 960 data were collected from the "Ilovica" treatment plant in Bosilovo, Republic of Macedonia, over of four years. This was done to provide a wide range of values for the variables and to cover different processing periods.

2 METHODOLOGY

2.1 Water Treatment Process

The water treatment plant "Ilovica" has the ability to treat 3840 cubic meters of water per day or a maximum of 44 litters per second. The treatment process is fully automated and controlled by a Programmable Logic Controller, the stages (phases) are presented in Fig. 1.

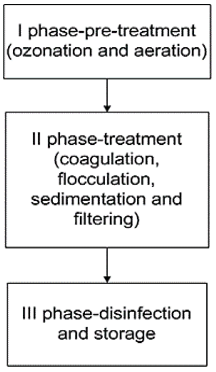
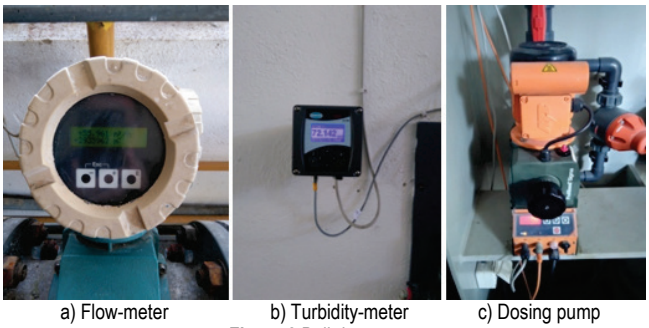


Figure 1 Different phasess of water treatment

In order to effectively control the process, it is important to measure and record the variables being monitored. The way these variables are measured is determined by the existing monitoring system. Some variables are measured in real time using sensors embedded in the system, while others are measured manually using portable measuring devices.



a) Flow-meter b) Turbidity-meter c) Dosing pump

Figure 2 Built-in sensors

2.2 Knowledge by Operators

The goal of using experiential knowledge to find a solution is to relate the values of the input variables to the values of the output variables [17]. The hidden knowledge, represented by historical data can be expressed as a vector (x_i, y_i), i = 1, 2, ..., n (n - number of pairs). Experienced operators can use their knowledge to determine processing rules and define the values of variables in groups, through linguistic expression. The number of selected if-then rules depends on the number of selected groups (subsets). This is shown in Fig. 3. Four distinct groups or subsets represent the qualitative definition of the range of values for the two input variables: small (C_1), medium-small (C_2), medium (C_3), big (C_4) for the flow of the raw water (x_1) and low (D_1), medium-low (D_2), medium (D_3), high (D_4) for the turbidity of the raw water (x_2). The optimal connection of input-output variables (flow and turbidity of raw water like input variables and amount of dosage of flocculent solution like output variable) in our case is set with "if-then" rules (Fig. 3). The calculations for each rule are done using Eq. (1) separately.

If x_1 is in A_i and x_2 is in B_j and i = 1, 2, ..., n and j = 1, 2, ..., n (n = 4) then,

$$y = x_1 B_{1(i,j)} + x_2 B_{2(i,j)} + B_{0(i,j)} \tag{1}$$

Where: A_i - groups (flow of raw water); B_j - groups (turbidity of raw water); x_1, x_2 - input variables (flow and turbidity of raw water); y - output variable (amount of dosage of flocculent solution); B_0, B_1, B_2 – coefficients; n - number of groups (subsets).

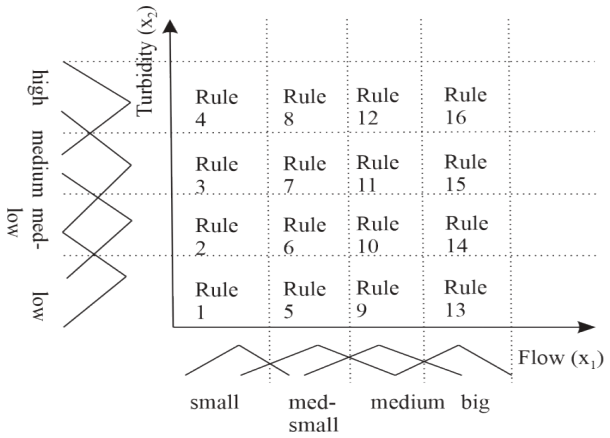


Figure 3 Selected "if-then" rules

Before creating a non-uniform network partition, the number of variables and groups (subsets) to be used are determined. Then, learning methods are used to refine and optimize the previous parameters [18]. This "fine-tuning" capability is used here, in order to determine the boundaries of each of the groups (subsets), as an intersection between two membership functions (Fig. 3). The boundaries of each group (subset) are used in dosing calculation program and are determined by arithmetic calculation using the values of the membership function parameters. Obtaining the optimized parameters was done by using an intelligent ANFIS (Adaptive Neuro Fuzzy Inference System) tool.

2.3 Data Preparation

If any of the variables listed in Tab. 1 change, it implies a change in the water treatment process in real time. In our case, water treatment is done by rapid treatment and pressure filtration, and as emphasized earlier, any change in the input variables directly affects the water treatment. The relationship between the input and output variables is shown by a functional dependency:

$$y = B_0 + x_1 B_1 + x_2 B_2 \tag{2}$$

Where: y - output variable; x_1, x_2 - input variables; B_0, B_1, B_2 - coefficients.

Table 1 Observed values of variables

	Flow	Turbidity	Dosing of flocculent
	m³/h	NTU	l/h
Maximum	161	99	130
Average	83	53	51
Minimum	29	6	16

The collected input-output pairs for a given criterion (a turbidity value of less than 0.100 NTU) it depends on the processing regime and the total volume in the treatment stage:

$$NS = \frac{VT}{RP} \tag{3}$$

Where: NS - number of samples in one observation; VT - total volume of phase treatment; RP - regime of processing.

Tab. 2 shows the water treatment stage, which consist of coagulation and flocculation, sedimentation and filtration volumes.

Table 2 Total volume of phase treatment

Steps of phase treatment	Volume (m³)
Coagulation and flocculation	130.4
Sedimentation	120.4
Filtering	26.7
Total:	277.5

2.4 ANFIS (Adaptive Neuro Fuzzy Inference System)

ANFIS (Adaptive Neuro Fuzzy Inference System), introduced by Jang, is a rule-based approximation tool [18]. A specific Sugeno type fuzzy system is presented, which is implemented in a five-layer feed-forward network structure [19]. ANFIS uses "if-then" rules to determine output based on input values [20]. ANFIS as an intelligent tool uses observed and historical data from operational lists to train. Such a tool can also be used in water treatment systems [21]. The ANFIS architecture we are using is a zero-order Sugeno fuzzy model. It includes two inputs, x_1 and x_2, and one output, y. In Fig. 4 is shown the general structure of ANFIS.

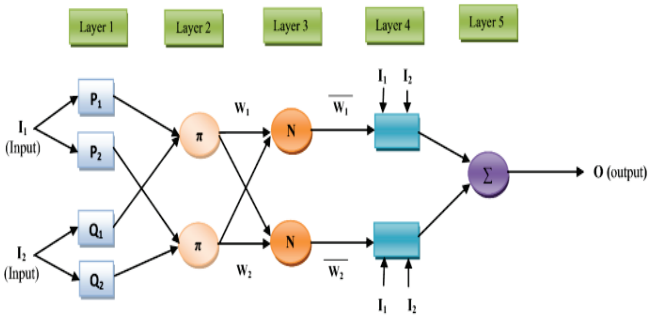


Figure 4 Structure of ANFIS (Adaptive Neuro Fuzzy Inference System)

The hybrid Neuro Fuzzy tool ANFIS optimizes the parameters. These optimized parameters are referred to as "previous" parameters, while the constant values for the output are known as "consequential" parameters. ANFIS learning involves a combination of back propagation and Least Squares Estimation (LSE) [20]. Using the Fuzzy Logic Toolbox, the optimized parameters were obtained. To create a fuzzy approximation system, a network partition is imitated. The values of "previous" parameters that define the triangular fuzzy membership functions, obtained after optimization and which are used in the model shown in Fig. 5, are shown in Tab. 3.

Table 3 Triangular fuzzy membership functions

Linguistic expression	Values of parameters
Small	[3.6; 4.78; 5.94]
Medium-small	[4.76; 5.94; 7.17]
Medium	[5.95; 7.09; 8.30]
Big	[7.13; 8.27; 9.45]
Low	[0.04; 4.92; 9.88]
Medium-low	[4.92; 9.89; 4.84]
Medium	[9.87; 14.84; 19.8]
High	[14.84; 19.8; 24.76]

2.5 Linear Regression

Linear regression is a method-tool used to identify the relationship between variables and the strength of that relationship. Dependence can be functional and statistical.

The relation represents functional dependence:

$$u = f(v) \tag{4}$$

The relationship is shown by the equation:

$$u = f(v) + e \tag{5}$$

Where: $f(v)$ is a function component, and e - error component.

Here, we will use a multiple linear regression model. This model is represented by a general formula:

$$u = q_0 + q_1 p_1 + q_2 p_2 + \ldots + q_i p_i + e \tag{6}$$

The variable u is dependent in the given relationship, p_i are independent variables, q_0, q_1, q_2, …, q_i are coefficients and e is stochastic variable (error).

The functional dependence between the variable u and variables p_i is represented as:

$$u = q_0 + q_1 p_1 + q_2 p_2 + \ldots + q_i p_i \tag{7}$$

A linear regression model shows how the variable u is related to the variables p_i, determined using samples (pairs). This model can be expressed as a system of n-relations.

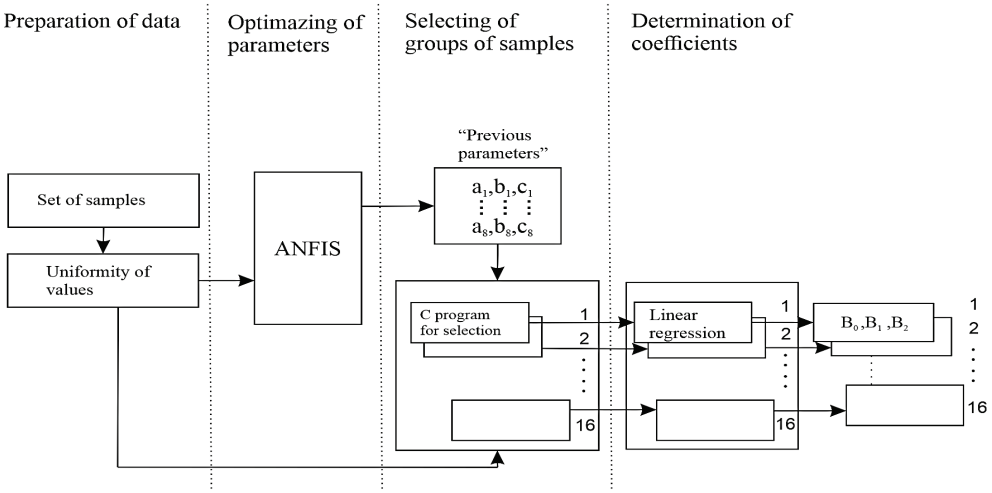


Figure 5 Path-procedure for obtaining the coefficients (ANFIS–Regression model)

In this research is used Regression in Excel to find the coefficients B_0, B_1 and B_2 used in Eq. (1). To determine the number of pairs for each rule, an algorithm-program was developed in the C programming language. Obtaining the coefficients applied in the dosing calculation program (Fig. 13) is show in the Fig. 5.

3 RESULTS

The performance and indication of accuracy of the model is expressed through Mean Absolute Percentage Error (MAPE) Eq. (8) and with coefficient of determination R^2 Eq. (9), using 240 samples (pairs). The results are shown in Tab. 4.

Table 4 Model performance

Model	*MAPE*	R^2
ANFIS-Regression	6.3	0.980

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|o_i - d_i|}{o_i}\times 100 \quad (8)$$

Where: *MAPE* - Mean Absolute Percentage Error; *o* - observed values; *d* - determined (predicted) values; *n* - number of samples.

$$R^2 = 1 - \left(\frac{\sum_{i=1}^{n}(o_i - d_i)^2}{\sum_{i=1}^{n} o_i^2}\right),\ i = 1, 2, ..., n \quad (9)$$

Where: R^2 - coefficient of determination; *o* - observed values; *d* - determined (predicted) values; *n* - number of samples.

The results of observed values of input and output variables shown in Figs. 6, 7, 8, and the determined (calculated) values of dosing of flocculent (0.15% solution of anionic flocculent) shown in Fig. 9. The calculation of dosing values of flocculent with Rule Viewer from MATLAB Fuzzy Logic Toolbox is shown in Fig. 10**.**

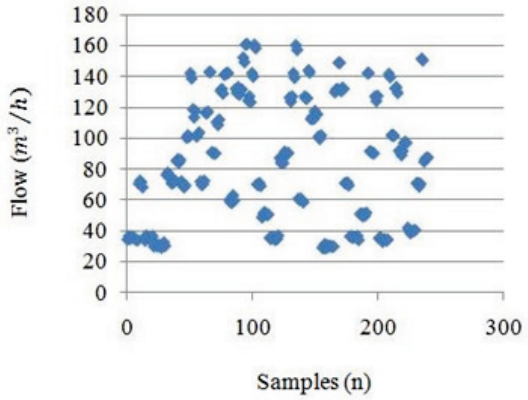


Figure 6 Observed values of flow of row water

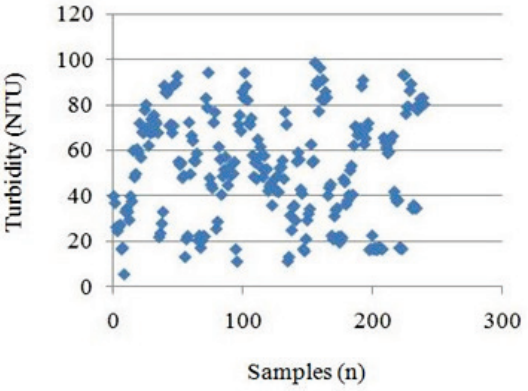


Figure 7 Observed values of turbidity of row water

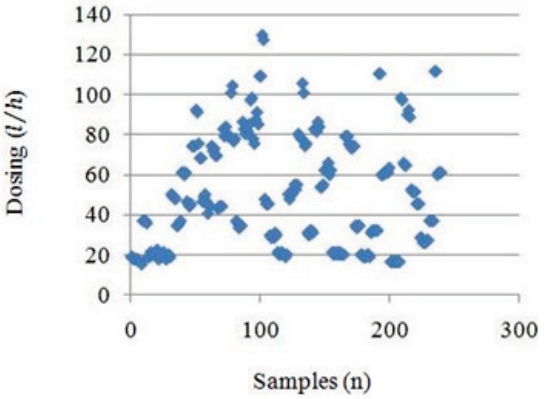


Figure 8 Observed dosing values of flocculent

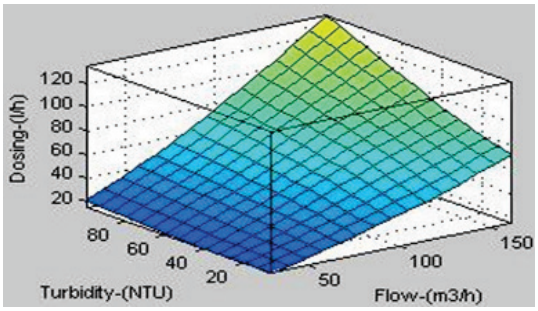


Figure 9 ANFIS surface (determined flocculent)

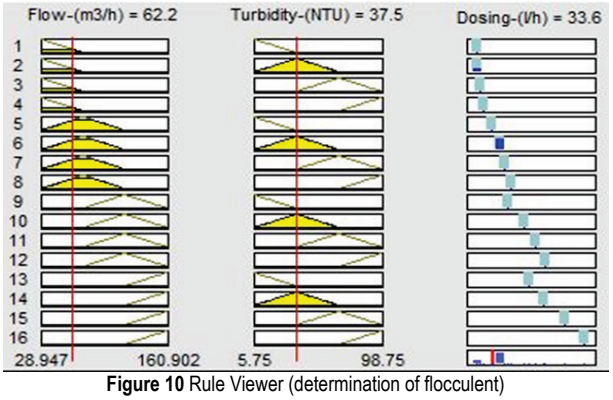


Figure 10 Rule Viewer (determination of flocculent)

4 CONCEPT AND GRAPHIC PRESENTATION

Fig. 11 shows the concept of incorporating a dose calculation program into the dosing system as a control unit. Here the concept for dose determination, which graphically is presented trough diagram in Fig. 12.

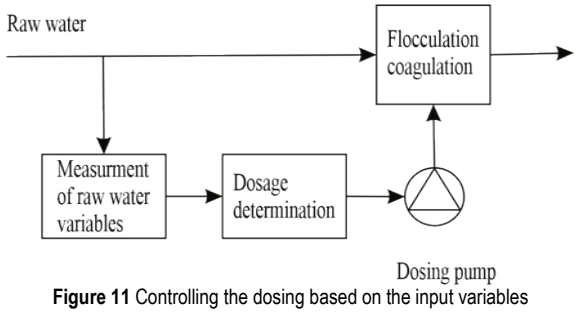


Figure 11 Controlling the dosing based on the input variables

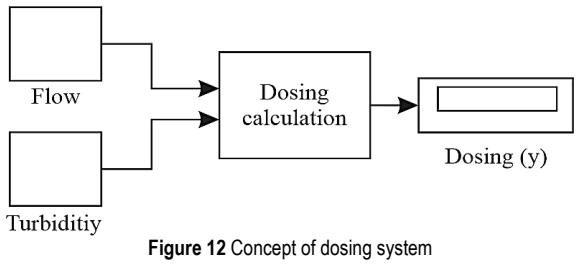


Figure 12 Concept of dosing system

```
IF x1 <= 5.35 AND x2 <= 7.40 THEN
    y := (x2 * 0.066) - (x1 * 0.009) + 5.301;
ELSIF x1 <= 5.35 AND x2 > 7.40 AND x2 <= 12.35 THEN
    y := (x1 * 0.649) + (x2 * 0.064) + 2.076;
 ELSIF x1 <= 5.35 AND x2 > 12.35 AND x2 <= 17.32 THEN
    y := (x1 * 2.455) + (x2 * 0.080) - 7.074;
ELSIF x1 <= 5.35 AND x2 > 17.32 THEN
    y := (x1 * 2.591) + (x2 * 0.068) - 7.582;
ELSIF x1 > 5.35 AND x1 <= 6.53 AND x2 <= 7.40 THEN
    y := (x1 * 1.988) + (x2 * 0.280) - 6.300;
 ELSIF x1 > 5.35 AND x1 <= 6.53 AND x2 > 7.40 AND x2 <= 12.35 THEN
    y := (x1 * 2.335) + (x2 * 0.172) - 7.556;
ELSIF x1 > 5.35 AND x1 <= 6.53 AND x2 > 12.35 AND x2 <= 17.32 THEN
    y := (x1 * 2.706) + (x2 * 0.142) - 9.428;
 ELSIF x1 > 5.35 AND x1 <= 6.53 AND x2 > 17.32 THEN
    y := (x1 * 1.911) + (x2 * 0.111) - 3.832;
ELSIF x1 > 6.53 AND x1 <= 7.70 AND x2 <= 7.40 THEN
    y := (x1 * 2.085) + (x2 * 0.153) - 6.157
ELSIF x1 > 6.53 AND x1 <= 7.70 AND x2 > 7.40 AND x2 <= 12.35 THEN
    y := (x1 * 2.667) + (x2 * 0.324) - 11.548;
ELSIF x1 > 6.53 AND x1 <= 7.70 AND x2 > 12.35 AND x2 <= 17.32 THEN
    y := (x1 * 2.932) + (x2 * 0.297) - 13.172;
ELSIF x1 > 6.53 AND x1 <= 7.70 AND x2 > 17.32 THEN
    y := (x1 * 3.260) + (x2 * 0.004) - 10.366;
 ELSIF x1 > 7.70 AND x2 <= 7.40 THEN
    y := (x1 * 2.251) + (x2 * 0.356) - 9.079;
ELSIF x1 > 7.70 AND x2 > 7.40 AND x2 <= 12.35 THEN
    y := (x1 * 2.147) + (x2 * 0.435) - 8.750;
ELSIF x1 > 7.70 AND x2 > 12.35 AND x2 <= 17.32 THEN
    y := (x1 * 2.217) + (x2 * 0.356) - 8.410
 ELSE
y := (x1 * 3.971) + (x2 * 0.066) - 16.981;

END_IF;
```

Figure 13 Dosing calculation program "if-then" form

5 CONCLUSION

The possibility of "flexible" control and application of the built-in equipment in a modern processing system has a crucial role in the fulfilment of the final task. This research aims to apply the world's scientific tendencies in the use of intelligent tools (methods) in the direction of advancement and improvement of the operation as a whole through optimization and modelling by applying the numerous data in the form of input-output variables. In this paper is presents a concept for determining the dose of flocculent in a real water treatment system. Since there is no generally accepted mathematical correlation, i.e. the characteristics of coagulation-flocculation in water treatment systems are different; an important aspect in this research was the establishment of functional dependence between some of the "input" parameters (variables) with the chemical-physical process of coagulation-flocculation. In this paper, a dosing calculation program ("ELSIF"-statement), is developed in PLC platform "Codesys" in Structured Text (ST), with the possibility of applicability in a logic controller (Fig. 13). The model was obtained by integrating hybrid Neuro Fuzzy tool ANFIS and Regression. The statistical indication, Mean Absolute Error (MAPE) and coefficient of determination R^2 was used in order to determine the performance of the applied model.

The disadvantage of this model is the exponential increase in processing rules with the addition of variables.

Advantages of this research, as well as future recommendations are:

A simplified "if-then" dosing program for use in logic controllers.

Possibility of implementation of other input-output parameters (variables) for water treatment such as conductivity, acidity (pH), temperature, color, UV absorption, i.e. doses of solutions for coagulation-flocculation and disinfection.

Measurements performed in real time by all embedded processes meters-transmitters, and from which operational data is collected (recorded), can be applied in order to obtain conceptual models such as dosing systems.

The implementation of other process measurements in real time, such as conductivity measurement in order to manage coagulant dosing, as well as research on the replacement of coagulants using "electrocoagulation" which is a coagulation process without the use of chemicals for coagulation.

6 REFERENCES

[1] Pingale, R., & Shinde, S. N. (2021). Fuzzy logic approach for routing in Internet of Things network. *Tehnički glasnik*, *15*(1), 18–24. https://doi.org/10.31803/tg-20210204155601

[2] Svalina, I., Galzina, V., Lujic, R., & Simunovic, G. (2013). An adaptive network-based inference system (ANFIS) for the forecasting: The case of close price indices. *Expert Systems with Applications*, *40*(15), 6055–6063. https://doi.org/10.1016/j.eswa.2013.05.029

[3] Valcic, M., Antonic, R., & Tomas, V. (2011). ANFIS based model for ship speed prediction. *Brodogradnja*, *62*(4), 373–382. https://doi.org/10.21278/brod

[4] Dimanoski, K., Vrtanoski, G., & Stojic, G. (2017). Fuzzy logic model for dimensioning of transport capacities in railway freight transport. *Mechanical Engineering Scientific Journal*, *35*(2), 105–114.

[5] Yetkin, M. E., Simsir, F., & Ozfirat, M. K. (2016). A fuzzy approach to selecting roof supports in longwall mining. *South African Journal of Industrial Engineering*, *27*(1), 162–177. https://doi.org/10.7166/27-1-1366

[6] Kahraman, C., Cebeci, U., & Ulukan, Z. (2003). Multi-criteria supplier selection using fuzzy AHP. *Logistics Information Management*, *16*(6), 382–394. https://doi.org/10.1108/09576050310503367

[7] Kahraman, C., Cebeci, U., & Ruan, D. (2004). Multi-attribute comparison of catering service companies using fuzzy AHP: The case of Turkey. *International Journal of Production Economics*, *87*(2), 171–184. https://doi.org/10.1016/S0925-5273(03)00099-9

[8] Kahraman, C., Cevik, S., & Oztaysi, B. (2015). Fuzzy multicriteria decision-making: A literature review. *International Journal of Computational Intelligence Systems*, *8*(4), 637–666. https://doi.org/10.1080/18756891.2015.1046325

[9] Ozifrat, P. M. (2020). A fuzzy event tree methodology modified to select and evaluate suppliers. *South African Journal of Industrial Engineering*, *31*(1). https://doi.org/10.7166/31-1-2027

[10] Wan, J.-Q., Huang, M.-Z., Ma, Y.-W., Guo, W. J., Wang, Y., & Zhang, H.-P. (2010). Control of coagulation process in a

paper–mill wastewater treatment process using a fuzzy neural network. *Chemical and Biochemical Engineering Quarterly*, *24*, 425–435. https://hrcak.srce.hr/62712

[11] Heddam, S., Bermad, A., & Dechemi, N. (2012). ANFIS-based modelling for coagulant dosage in drinking water treatment plant: A case study. *Environmental Monitoring and Assessment*, *184*, 1953–1971. https://doi.org/10.1007/s10661-011-2091-x

[12] Trinh, T. K., & Skang, L. (2010). Response surface methodological approach to optimize the coagulation-flocculation process in drinking water treatment. *Chemical Engineering Research and Design*, *89*, 1126–1135. https://doi.org/10.1016/j.cherd.2010.12.004

[13] Juntunen, P., Liukkonen, M., Lehtola, M., & Hiltunen, Y. (2012). Dynamics of sensors for detecting factors affecting turbidity in drinking water. *Journal of Hydroinformatics*, *15*(2), 416–426. https://doi.org/10.2166/hydro.2012.052

[14] Boruszenko, P., Haley, J., & Thompson, K. C. (2015). On-line analysis: A water company perspective. https://www.rsc.org/images/JohnHaleyYWS5thNovCambridge_tcm18-143639.pdf

[15] Naidoo, P., & Van der Walt, J. J. (2013). Artificial neural networks as chemical dosing budgeting tool. https://hdl.handle.net/10520/EJC143029

[16] Cooperative Research Centre (CRC) for Water Quality and Treatment. (2015). *Natural organic matter – understanding and controlling the impact on water quality and water-treatment processes*. http://www.waterra.com.au/publications/document-search/

[17] Kasabov, N. (1996). *Foundation of neural networks, fuzzy systems and knowledge engineering*. The MIT Press. https://doi.org/10.7551/mitpress/3071.001.0001

[18] Yen, J., & Langari, R. (2013). *Fuzzy logic intelligence, control and information*. Pearson.

[19] Jang, J.-S. R. (1993). ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, *23*(3), 665–685. https://doi.org/10.1109/21.256541

[20] Gorzalczany, M. B. (2002). *Computational intelligent systems and applications: Neuro fuzzy and fuzzy neural synergisms*. Physica-Verlag.

[21] Wu, G. D., & Lo, S. L. (2008). Predicting real-time coagulant dosage in water treatment by artificial neural networks and adaptive network based fuzzy inference system. *Engineering Applications of Artificial Intelligence*, *21*, 1189–1195. https://doi.org/10.1016/j.engappai.2008.03.015

Authors' contacts:

Mirche Stojkov, MSc
(Corresponding author)
Public Utility Service Company "Ograzhden",
s. Bosilovo No. 257, 2431 Bosilovo,
Municipality of Bosilovo, Republic of North Macedonia
mircestojkov@hotmail.com

Zoran Pandilov, PhD, Full Professor
Institute of Production Engineering and Management,
Faculty of Mechanical Engineering,
Ss. Cyril and Methodius University in Skopje,
P. O. Box 464, 1001 Skopje, Republic of North Macedonia
zoran.pandilov@mf.edu.mk

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240410054928

Original scientific paper
Received: 2024-04-10, Accepted: 2024-11-13

Investigate Unsolicited Traffic on IoT Devices Using Machine Learning

Badeea Al Sukhni, Bhabendu Kumar Mohanta*, Debnath Bhattacharyya, Tai-hoon Kim*

Abstract: The significant growth and use of IoT devices, especially in smart homes, increase the risk of cyber-attacks compromising these devices and their services, resulting in data breaches and privacy invasion. Machine learning technology has been considered a sufficient security solution in IoT due to its effectiveness in detecting unsolicited traffic on an IoT network, ensuring the confidentiality, integrity, and availability of IoT devices, and, most importantly, protecting the users' privacy. In this research, different classification-based machine learning algorithms were leveraged to detect and classify different types of network traffic on the IoT network. At the same time, the effectiveness of these classification machine learning algorithms, Decision Tree (J48), Bayes Net, and Naive Bayes were conducted on our smart home dataset. Before the implementation of this algorithm, a man-in-the-middle attack was introduced on the IoT network while the network traffic was equally captured in the process. Overall, the algorithms successfully evaluated the captured network traffic, detected the introduced MITM attack on IoT devices, and classified the traffic into solicited and unsolicited traffic. According to the findings, the Naive Bayes algorithm outperformed the others with an accuracy of 99.95%.

Keywords: Arp Poisoning; Internet of Things (IoT); Machine Learning; MITM attack; Security; Smart home

1 INTRODUCTION

Internet of Things (IoT) refers to connecting and remotely managing the users' smart devices to each other and the Internet. Every device is embedded with storage, communication, and computing technologies [1]. Before 1990, the world had the internet, but without things. This indicates that there weren't intelligent cameras, home assistants, or even smart lights until the first IoT device, John Romkey and Simon Hackett, developed the internet toaster [2]. They used the TCP: Transmission Control Protocol/ Internet Protocol (TCP/IP) networking to connect the toaster to the internet. They also utilized the Simple Networking Management Protocol (SNMP) to control the power applied to the toast. In 1999, Kevin Ashton [3] coined the term IoT; he proposed a method to connect smart devices to the internet using primary enabling technologies such as Radio Frequency Identification (RFID) and Wireless Sensor Network (WSN) [4]. Since then, IoT has been a significant research topic in limitless areas, and the usage of IoT devices has also increased significantly. The International Data Corporation (IDC) predicts that by 2025, the data generated from the connected 55.7 billion IoT devices will be 73.1 Zettabytes (ZB), as these devices can monitor and gather information from their surroundings to perform complex tasks that benefit humanity. In other words, the IoT is the reason behind advanced manufacturing, smart homes, smart cities, automatic car tracking, modern surveillance, smart lighting, transportation, agriculture, and health care systems [5]. Many smart home products are available but unused due to various security and privacy challenges [6].

The plethora of usage of IoT devices raises the potential of cyber-attacks where the attackers compromise these devices and their services to achieve their demands. This could result in data leakage, as these devices exchange and store sensitive and private information. For instance, smartphones are used to browse the internet, pay bills, receive medical care, and sign in to various services, all of which require personal information. However, it may appear handy to use smartphones whenever and anywhere, with sensitive information stored on them, such as medical records, bank accounts, or passwords, which can be a security concern. According to a study conducted at the University of Pennsylvania, 68 percent of existing mobile devices are incapable of protecting against cyber-attacks [7]. An example of a cyber-attack that happened in 2011 was the DroidDream Malware, which infected 68 applications on the Android market. These apps were downloaded over 260K times in four days, and they used the Android Debug Bridge vulnerability on root phones and sent unauthorized premium-rate SMS messages late at night [7]. Since the network layer is the backbone for connecting smart devices to the internet, it could pose more security threats than other layers that compromise those devices. Some threats on this layer are Man-In-The-Middle attack (MITM), sniffing attack, Denial of Service (DoS), and network intrusion attacks. The most common security, privacy, and cybersecurity-related issues at this layer that were identified by [8-10] include:

- Heterogeneity: The different technologies and protocols [11] the devices use to communicate in the IoT network make security and network management more difficult to maintain.
- Scalability issues: Due to the massive number of IoT devices connected to the network, these devices can enter and leave anytime. This introduces new challenges to the IoT network, such as congestion and lack of authentication.
- Data disclosure: cyber-attackers may use social engineering techniques that may jeopardize sensitive data and result in stealing them from the network. Also, they can deploy different data recovery methods to extract critical information from the nodes. This leads to sufficient security approaches such as Machine Learning (ML), Intrusion Detection Systems (IDS), honeypots, and firewalls for preventing and detecting cyber-attacks, ensuring confidentiality, integrity, and availability of IoT devices, and finally, protecting the users' privacy invasion.

1.1 Research Aims and Objectives

This research aims to evaluate the effectiveness of the Decision Tree (J48), Bayes Net, and Naive Bayes algorithms in detecting MITM attacks on IoT devices and classifying the network traffic into solicited and unsolicited traffic. In this research, the Xiaomi Redmi Note 9S device was used as an example of IoT devices due to the remarkable growth of using smartphones in smart homes, as smartphones contain various kinds of sensors that store a large amount of sensitive information, enabling mobile applications to be used in a variety of IoT areas [12].

The objective is to conduct in-depth research on the Decision Tree (J48), Bayes Net, and Naive Bayes algorithms to understand them better and determine their significance in securing the Xiaomi device and differentiating between normal and abnormal network traffic. Also, to perform an MITM attack, specifically an ARP poisoning attack on the Xiaomi Redmi Note 9S device, because MITM attacks are one of the most common threats that compromise IoT devices and result in stealing sensitive information about users. Then, to capture the network traffic of the home network to which the Redmi Note 9S device is connected. Besides, to perform a critical analysis of the result produced by the Decision Tree (J48), Bayes Net, and Naive Bayes algorithms for detecting the unsolicited network traffic on the Redmi Note 9S device.

The rest of this paper is organized as follows: In Section 2, data collection, test environment, and system design with attack model are explained. Section 3 presents the implementation details. Section 4 discusses the experiment's results. The paper has been concluded, and future scopes are indicated in Section 5.

2 METHODOLOGY AND SYSTEM DESIGN

This section discusses the methodology of the proposed approach, which focuses on data collection techniques. The questionnaire and interview were two data collection methods employed in this research. They were used to collect opinions from home users regarding security breaches on IoT devices that result in privacy invasion. This section covered the design of the proposed system, the project test environment, which includes hardware and software requirements, and the network setup.

2.1 Data Collection

The first data collection stage involved interviewing home users of IoT devices to get their perspectives on privacy invasion concerns. The next stage involved developing a questionnaire based on the responses obtained from home users. These approaches have been used to determine their opinions regarding security breaches on IoT devices and to understand how they would like their IoT devices to be secured using an effective Naive Bayes Model. The rationale for using the questionnaire and interview approaches is that they are the most effective way to collect data from a large participant in a short period. They are also the most cost-effective method [13] since the data collection medium was online due to the COVID-19 pandemic.

The questionnaire comprises six questions using a close-ended approach in a multiple-choice form. The questionnaire's questions were selected and evaluated by me and IoT home users for the following reasons. This includes:

The first rationale is that we deeply considered instances where unauthorized/third-party attackers could access their home network without their consent. This is aligned with the primary objective of the entire research, which is to ensure that confidentiality is achieved in a network connection that interlinks IoT home users alongside the smart devices they use in their homes. However, after careful consideration with home users, we decided to use these questions as some had no prior knowledge of when a third party was accessing their wireless home network.

Secondly, some IoT home users fell prey to privacy invasion attacks where their secret data were exposed to attackers. For instance, one of the IoT home users we conversed with before developing the questionnaires aired their views on how their smart home camera was breached. This spurred our desire to incorporate specific questions about the privacy invasion experience for potential victims in my questionnaire.

Thirdly, after sharing their experiences, they suggested some defensive mechanisms to prevent them from the catastrophic privacy breaches they had encountered and other future attacks. This inspired us to search for significant questions about protecting their home network and IoT applications.

Overall, because some home users lack the technical know-how to utilize defensive mechanisms to secure their network from privacy concerns, they suggested automatic preventive and defensive measures, which allowed us to include this in the questionnaire.

Moreover, due to the rising number of IoT attacks in recent years and the current attacks within this global pandemic, we included a section relating to the collection of private data of IoT applications by IoT manufacturers. This is because we believe that effective and robust IoT data handling will equally mitigate the growth of IoT attacks, especially attacks dealing with privacy invasion.

Immediately after developing these electronic-based questionnaires, we directly sent them to IoT home users to get their consent on the issues to be addressed, how well they want them solved, and how it will help them stay protected and secured from privacy attacks. However, all responses on the questionnaires and feedback obtained were directly answered by only IoT home users.

2.2 Test Environment

This section discusses the design environment's software and hardware requirements and the network setup.

2.2.1 Hardware Requirements

The devices utilized in the project's implementation are listed in the table below. The Xiaomi Redmi Note 9S was a target device for an MITM attack. The Sony laptop was used for various purposes, including capturing network traffic with the Wireshark tool, introducing the ARP poisoning

attack on the smartphone with Kali Linux, and collecting the regular traffic from this device, as well as Alexa (Amazon Echo Dot 3rd gen) and the Apple HomePod mini.

Table 1 Devices used in the research implementation.

Device	Notes
Redmi Note 9S	Android OS, and 128 GB RAM
Sony laptop	Windows 10, 16 GB RAM, and 1TB SSD
Amazon Echo Dot 3rd Generation	Alexa integration
Apple HomePod mini	Siri integration

2.2.2 Software Requirements

This section explains the software that was utilized during the implementation phase.

- **Scanning Home Network Devices.** Advanced IP Scanner was used to find the devices connected to the home network and their IP and MAC addresses. Advanced IP Scanner is a free network scanning tool that retrieves the IP and MAC addresses of the devices connected to a LAN network to perform network analysis. Its user-friendly interface detects all network devices and provides remote control, including the ability to turn them on and off.
- **Capturing the Traffic.** This research used Wireshark, a joint open-source network analyzer software known as a network monitor or packet sniffer, to capture network traffic. The primary rationale for using this program is that it is a powerful tool for network traffic monitoring and troubleshooting, as it can scan any type of network, including Ethernets, Wi-Fi, Bluetooth, and even monitor mode [14]. It also captures raw bits and decodes them into a human-readable format for analysis.
- **Introducing ARP Poisoning Attack.** To implement an ARP poisoning attack on the home network, Kali Linux must be installed on Windows 10 using Oracle Virtualbox to act like a Hacking machine, as the virtual machine provides a realistic environment. After installing Kali Linux in the Virtualbox, the ARP table is needed to check the IP addresses and the associated MAC addresses of the default gateway and Kali Linux are working as expected. This can be done by entering arp –an in the command prompt on a Windows 10 machine. To implement an ARP poisoning attack on the home network, the Arpspoof tool must be installed on Kali Linux and the dsniff package, including the Arpspoof tool. The Arpspoof tool is very effective for sniffing network traffic in a home network, and it's simple to use because the coding is straightforward [15]. It also alters the flow of packets such that they pass via the attacker's machine. Besides, various attacks can be performed on the same platform because Arpspoof comes with the entire Dsniff package [15].
- **Deploying Decision Tree (J48), Bayes Net and Naive Bayes algorithms.** To evaluate the proposed solution, a Python script is needed to extract the features and convert the .pcap file to .csv. At this point, installing Python 3 on Kali Linux and the Tshark tool is required. Tshark is part of the Wireshark tool that makes the .pcap file more straightforward to understand, extracts the packets from the pcap file, and converts it to a vector space model. Besides, the Weka tool must also be installed on Windows 10. Weka is a machine learning toolkit used to implement the Decision Tree (J48), Bayes Net, and Naive Bayes algorithms to easily apply to the research dataset, as the algorithms are available as an open-source through the Weka interface. This tool provides a graphical user interface for various machine-learning algorithms for data analysis and predictive modeling [16]. Also, it allows preprocessing of a dataset, feeding it directly into machine learning algorithms, and examining the final classifier and its performance without writing any code [16].

2.2.3 Network Setup

As mentioned, Kali Linux runs on Windows 10 devices using Oracle Virtualbox. Virtual machines have different networking modes, such as host-only adapters, internal networks, bridged adapters, and NAT networks. A bridged adapter network mode was chosen since the Xiaomi Redmi Note 9S device and Kali Linux must be connected to the same internal network environment to create the ARP poisoning attack. That's because bridge mode allows Kali Linux and the Xiaomi Redmi Note 9S device to access each other in the network. However, NAT mode allows one-way access. Kali Linux can access them through the network, but the Xiaomi Redmi Note 9S device cannot access it. Also, they cannot access each other in the Host-only adapter and internal mode because they are not in the same network.

2.2.4 System Design

As discussed earlier, Decision Tree (J48), Bayes Net, and Naive Bayes algorithms were employed in this research to evaluate its effectiveness in investigating unsolicited traffic, classifying it as normal and abnormal traffic, and protecting IoT devices from privacy invasion. Also, the IoT devices used in this work were the Xiaomi Redmi Note 9S, Sony laptop, Alexa, and HomePod mini device to test how to capture the traffic. The abnormal traffic was collected from Xiaomi devices, and the normal traffic was gathered from Sony laptops, Alexa, and HomePod mini devices. The design of the proposed system combines both the flowchart and pseudocode.

2.2.4.1 Flowchart

A flowchart is a diagram that is used to represent an algorithm. It shows the steps of writing the algorithm in appropriate shapes and connecting them with arrows to direct the flow of the processes. The flowchart shown in Figure 1 starts with the data collection plan for capturing the network traffic, and it ends with evaluating the performance of the Decision Tree (J48), Bayes Net, and Naive Bayes algorithms that were applied to the vector space model. After the network traffic has been collected and captured, the next two steps are utilized to extract the features and build a vector space model, which is then used to represent the network traffic in a format suitable for the machine learning algorithm. The following step deals with applying the feature selection method and reducing the extracted features, then using the three classifiers on the vector space model to evaluate the efficacy of all the extracted features or applying the three classifiers directly with the feature selection process. The final step deals with assessing the performance

of the algorithms in detecting the attack and classifying the traffic.

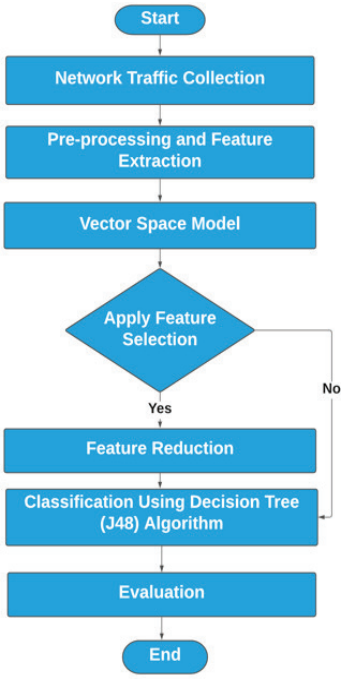


Figure 1 Flowchart of the proposed system for investigating unsolicited traffic

- **Network Traffic Collection.** The Xiaomi Redmi Note 9S, Alexa, Sony laptop, and HomePod were connected to the home network to collect network traffic. Wireshark's packet sniffer tool was used to capture the raw data for an extended period while capturing the traffic. An ARP poisoning attack was performed on the Xiaomi device, as shown in Fig. 2.



Figure 2 Network traffic collection

- **Pre-processing and Feature Extraction.** After collecting the home network traffic, the next step is to preprocess the data and extract the features. This step removes unnecessary features from the raw data to create patterns involving legitimate and suspicious activities [17].
- **Vector Space Model.** The following process applies the data to the vector space model. A vector space model is an n-dimensional vector of unique features that represents the data geometrically so that it fits into machine learning algorithms.
- **Feature Selection.** The feature selection process improves the effectiveness of machine learning algorithms and data classification performance [18]. Furthermore, since the dataset contains many non-essential features, the feature selection procedure can help reduce the number of extracted features by eliminating irrelevant features [18]. The feature selection process can also improve learning accuracy and speed up the modeling process [18].
- **Classification Using Decision Tree Algorithm.** The next step is to apply the open-source Decision Tree classifier using the Weka interface to the prepared dataset to detect the ARP poisoning attack on the home network and distinguish traffic between normal and abnormal behavior.
- **Evaluation.** The evaluation step calculates the proposed machine learning algorithm's accuracy in detecting attacks and classifying the traffic.

2.2.4.2 Pseudocode

A pseudocode was used to evaluate an algorithm to detect unsolicited traffic on the IoT network. Pseudocode is a text-based approach that helps design and develop algorithms. It uses informal English that does not use any particular programming language to describe the steps of an algorithm in a way that anyone with basic programming knowledge can understand. Then, a suitable programming language can be used to code the algorithm. The following pseudocodes in algorithms one and two demonstrate the ARP poisoning attack and the proposed approach to detect unsolicited traffic.

The above pseudocode in algorithm 1 demonstrates the ARP poisoning attack on the Xiaomi device. The first step after connecting the target and the attacker machine on the same network is to collect the traffic for further analysis. Then, IP forwarding is enabled to route the packets through the attacker's machine, preventing the target system from losing internet connectivity. The next step is to configure ARP poisoning. This can be done by launching the first mentioned command, which is used to ARP spoof the target device and inform it that the attacker's MAC address is the router's IP address. However, the following command in the configuration is used to ARP spoof the router and inform it that the attacker's MAC address is the router's IP address. The pseudocode's final step is to browse the Xiaomi smartphone to sniff the user's personal information.

As presented in the above pseudocode in algorithm 2, the input is the captured network traffic after introducing the MITM attack on the IoT network. The expected output is classifying the traffic into normal and abnormal traffic. The first step in the procedure is to create a Python script to extract the features from the captured network traffic. Then, filter the data from the .pcap file to store the ARP, DNS, and HTTP packets on the Redmi Note 9S IP address and consider them Abnormal_Traffic. Consider the TCP on the Sony laptop, Alexa, and HomePod IP address as Normal_Traffic.

After pre-processing the data, the next step is creating the vector space model and writing the string "label" and 33 features to the header of the .csv file. After preparing the headers of the .csv file, the next step is to filter all the .pcap file packets and store them into two .pcap files with the label name, as this step will reduce the raw data into relevant

packets. The next step is writing the content of these files into the .csv file to apply it to the Decision Tree (J48), Bayes Net, and Naive Bayes algorithms to detect the unsolicited traffic and classify the traffic.

3 IMPLEMENTATION

This section discusses the implementation of the proposed solution on the Xiaomi Redmi Note 9S device. The Xiaomi Redmi Note 9S smartphone, Sony laptop, Alexa, and Apple HomePod device were used to demonstrate a real-world IoT network. The Xiaomi smartphone was utilized as a target device for an MITM attack, specifically ARP poisoning, to capture the abnormal traffic. The average traffic was collected using a Sony laptop, Alexa, and the HomePod. The test environment, including the hardware, software, and network configurations, was also handled. Following the collection of network traffic, a Python script was created to extract features and build a vector space model. The process ended by preparing a dataset with 33 attributes and storing 13980 instances; 12240 instances are normal traffic, and 1740 instances are abnormal traffic. The implementation was concluded using the Weka interface to apply the three classifiers to the dataset.

3.1 Network Traffic Collection

Collecting sufficient amounts of network traffic in both standard and abnormal states is needed to distinguish the IoT network between normal and abnormal. We followed an approach proposed by [19] to collect the usual traffic. Using the Wireshark tool, we captured the regular traffic from seven IoT devices installed in a private home network for five days. In this research, we connected a Sony laptop, Alexa, and HomePod to the home router and captured the traffic for an extended period using Wireshark. Then, we connected the Xiaomi Redmi Note 9S device to collect the abnormal traffic to the network. We also used Virtualbox to launch Kali Linux and selected the bridge network mode. We also used the Advanced IP Scanner tool to find the IP and MAC addresses of the devices connected to the home network. Tab. 2 lists the IP and MAC addresses of each device.

Table 2 The IP and MAC addresses of the devices connected to the home network

Name	IP	MAC address
Home network router	192.168.8.1	94:E9:EE:8D:81:91
RedmiNote9S	192.168.8.100	18:87:40:EF:19:B1
Kali Linux	192.168.8.157	08:00:27:5C:65:26
Sony laptop	192.168.8.144	00:23:14:52:66:58
HomePod	192.168.8.152	4C:20:B8:DB:B2:8C
Alexa	192.168.8.111	08:84:9D:D5:C7:76

3.2 Pre-processing and Feature Extraction

Before pre-processing the captured PCAP file, we created a Python script to filter the collected traffic to have only the TCP, DNS, HTTP, and ARP packets. The TCP protocol has been chosen to represent the Normal traffic because this traffic has been collected from the Sony laptop's IP address, Alexa's IP address, and HomePod's IP address. On the other hand, the ARP, DNS, and HTTP packets were chosen to represent the Abnormal traffic, as security breaches and data leakage can occur if one or more of these protocols are successfully exploited [17]. Because the ARP protocol maps logical (IP) addresses to physical (MAC) addresses and launching an ARP poisoning attack allows the capture of the DNS packets that were requested from the user as the DNS helps internet users access their requested websites using domain names because the role of the DNS is to translate domain names into IPs [17]. Also, The HTTP protocol's purpose is to exchange hypertext, which is a textual format with hyperlinks [17]. From here, two classes were used in the research: Normal_Traffic and Abnormal_Traffic, based on the IP address and the protocol.

To keep it simple and specific to the problem, we focused on extracting 33 features from the ARP, DNS, HTTP, IP, and TCP headers to represent the raw data in the vector space model.

Because the .pcap file is vast and contains a lot of packets and data, as the list of the required IPs and protocols was discussed, we followed the [20] approach to filtering the .pcap file into two individual pcap files (with the same names as the selected labels), each one containing the data of a particular class. This process helps convert the .pcap to a CSV file with the extracted features quickly and easily.

Algorithm 1: Introducing ARP poisoning attack

Input: Connecting Xiaomi Redmi Note 9S and Kali Linux to the same network
Output: Introducing ARP poisoning attack on Xiaomi Redmi Note 9S
Begin:
1. Capture the traffic.
2. Configure the Kali Linux machine for IP forwarding.
   echo 1 > /proc/sys/net/ipv4/ip_forward
3. Configure ARP poisoning
   arp spoof –i eth0 –t target machine IP address –r router IP address
   arp spoof –i eth0 –t router IP address -r target machine IP address
4. Browse Xiaomi Redmi Note 9S

End

Algorithm 2: Detecting the unsolicited traffic

Input: The captured network traffic
Output: Classifying the traffic into normal and abnormal traffic
Begin:
1. Create a Python script to extract the features.
2. Filter the data from the .pcap file.
3. if (arp.dst.proto_ipv4 ==Redmi Note 9S IP address) || (dns && (ip.dst==Redmi Note 9S IP address)) || (http && (ip.dst== Redmi Note 9S IP address)
4. [Abnormal_Traffic]= these protocols data.
5. if (TCP && (ip.src==Sony laptop IP address) || (ip.src==Alexa IP address) || (ip.src==Apple HomePod IP address))
6. [Normal_Traffic] = this protocol data.
7. Create a vector space model file to store the features and labels.

8. Write the string "label" and 33 features to the header of the .csv file.
9. Filter all the .pcap file packets and store them into two .pcap
   files with the label name.
10. Write the content of these files in the .csv file.
11. Use the Decision Tree J48 algorithm
12. Calculate the accuracy of the Decision Tree J48 model

End

3.3 Vector Space Model

As mentioned earlier, a vector space model is necessary to represent the data geometrically so that it can easily fit into machine learning algorithms. Eq. (1) below shows how to define the vector space model; this equation is written by [17].

$$I_{1:N} = \begin{bmatrix} f_{11} & f_{12} & \dots \\ f_{21} & f_{22} & f_{ij} \end{bmatrix} \quad Y_{1:N} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}. \tag{1}$$

From Eq. (1), I is network traffic observations, Y is the classes (c), which are Normal_Traffic and Abnormal_Traffic, N is the number of observations, and f is the features, which are the fields of TCP, IP, and DNS, HTTP, and ARP headers to write the Python script to extract the features from the Pcap file and convert it to a CSV file, which is aligned with the proposed system.

It was necessary to import the required libraries into the script to extract helpful information from files and directory names, such as os and glob; os and glob libraries are used to access files and directories. Then, a dictionary called ip_filter was defined to filter the data from the .pcap file. Abnormal_Traffic and Normal_Traffic are the keys of the filtering string. Also, they represent the classes of the dataset. (arp. dst.proto_ipv4 == 192.168.8.100) || (dns (ip.dst==192.168.8.100)) || (http (ip.dst==192.168.8.100)) means to filter the ARP, DNS, and HTTP packets from 192.168.1.100, which is the IP address of Redmi Note 9s device and consider them as the abnormal traffic. On the other hand, (TCP (ip.src==192.168.8.144) || (ip.src==192.168.8.111) || (ip.src==192.168.8.152)) means to filter the TCP packets from 192.168.1. 144 (the IP address of the Sony laptop), 192.168.1. 111 (which is Alexa's IP address) and 192.168.1. 152 (the IP address of the HomePod) represents the usual traffic.

The next step is to create a CSV file to write the string "label" and 33 features to the header of the .csv file. Next, filter all the .pcap file packets and store them into two .pcap files with the label name. To do that, glob. The glob function from the glob library was used to process the pcap file in the NetworkTraffic folder. Also, the T-shark was used to write the packets matching the particular filter, as these filtered pcap files will be stored in the Filtered_NetworkTraffic folder. The next step is to write the content of new pcap files into the CSV file. Again glob. Glob function was used to process the pcap files in the Filtered_NetworkTraffic folder. Also, the T-shark command is used to extract features from the new pcap files. After this process, the os. The open function was used to execute the T-shark command and save the result to the features variable. This results in preparing the dataset to be applied to the three classifiers and storing 13980 instances (12240 instances as regular traffic and 1740 instances as abnormal traffic) with 33 attributes in addition to the "Label" header in the CSV file.

4 RESULTS ANALYSIS

This section analyzes the questionnaire results and evaluates the performance of the Decision Tree algorithm in detecting unsolicited traffic and classifying it using the Weka interface.

4.1 Questionnaire Results

The questionnaire consists of six questions, and we received responses from 100 home users who completed it. It begins with how they would know if a third party obtained unauthorized access to their home network. 64% said they would get help from a network technician, 32% said they wouldn't know, and 4% said they would find out from a friend. While 81.8% had experienced a data breach or had their network attacked. 85.9% of them realized their IoT network had been hacked with the help of a technical expert, 12.9% of them from a friend, and 1.2% from their parents. Besides, 84% of them want their IoT network to be protected from cyber-attacks via a system that automatically detects attacks and classifies them as normal or abnormal, 15% want a system that detects the attack and shuts down the home network, and 1% want a system that detects the attack. Finally, 59% said IoT device manufacturers should conform to stricter privacy regulations. As a result of this questionnaire, the requirements of the home users were better understood and aligned with the proposed solution.

4.2 Classification Accuracy

The Weka interface was used for the classification process. Weka provides several services, such as data pre-processing, classification, clustering, association, visualization, and feature selection [18]. As stated earlier, WEKA also provides various open-source algorithms developed using Java programming to conduct data mining processes [18]. These algorithms can be used directly on the dataset or developed and changed using Java code.

In this research, we utilized the open-source Decision Tree (J48), Bayes Net, and Naive Bayes algorithms to evaluate their effectiveness in detecting and classifying unsolicited traffic that could compromise network confidentiality. The classification process has two stages: classification before and after feature selection.

4.2.1 Classification before Feature Selection

Before applying feature selection to the dataset, all 33 features were used to evaluate the accuracy of the three classifiers. Four metrics, accuracy, precision, recall, and F1 (F-measure), were used to evaluate the performance. A confusion matrix was used to introduce them, as it shows how many elements from a class are correctly and incorrectly classified. Tab. 3 demonstrates the confusion matrix.

Table 3 Confusion matrix

	Classified Positive	Classified Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

Where [21]:

- TP: the true positive is the actual average records correctly classified as benign behaviors.
- TN: the true negative indicates that abnormal records are correctly classified as attacks.
- FP: the false positive indicates that attack behaviors are classified incorrectly as benign.
- FN: the false negative indicates that benign records are incorrectly classified as attacks.

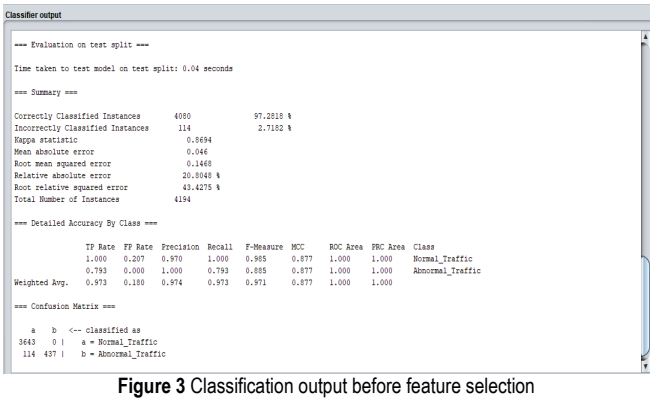


Figure 3 Classification output before feature selection

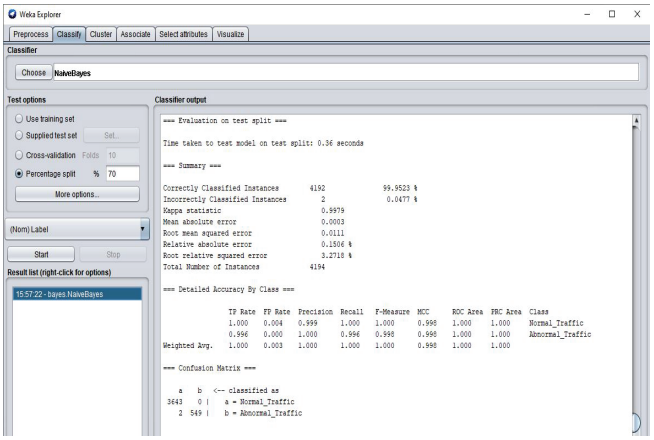


Figure 4 Naive Bayes result

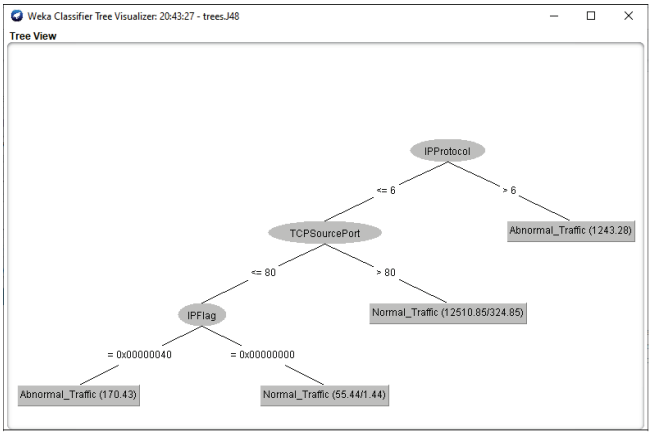


Figure 5 Visualization of the result of J48

Fig. 3 shows the classification results before feature selection. From the confusion matrix, it can be concluded that the total correctly detected traffic was 4080, with 3643 normal traffic and 437 abnormal traffic. However, 114 regular traffic was incorrectly detected as abnormal. Fig. 4 shows the result of the Naive Bayes algorithm using Weka.

Fig. 5 demonstrates the visualization of J48 results. It shows that the tree was 7 feet tall and had four leaves.

4.2.2 Classification after Feature Selection

In this research, the attribute evaluator, which is InfoGainAttributeEval, was employed with the ranker search method in the Weka interface. The InfoGainAttributeEval evaluates how each extracted feature decreases the input dataset's overall entropy [20]. A relevant attribute with the most significant information and hence reduces entropy the most could be prioritized above another using this strategy [20]. This technique was used to rank each attribute from the most significant to the least important.

In addition to the attribute evaluator and its search method, the attribute selection mode was also selected. In this research, a 10-fold cross-validation method was chosen for the dataset, as the 10-fold cross-validation process can effectively determine the value of the attribute subset. The 10-fold means the process was repeated 10 times, as the nine folds were used as training data, and one fold was used as testing data.

The results of applying feature selection showed that 17 of the features were irrelevant, and 16 features relating to IP and TCP headers were relevant to the problem. The IP checksum had the highest rank. Fig. 6 presents the attribute selection output.

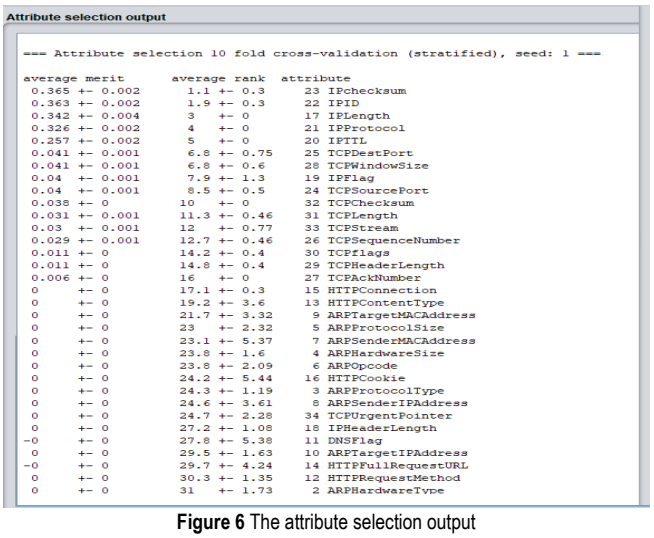


Figure 6 The attribute selection output

Table 4 The evaluation results of three algorithms for detecting ARP poisoning attack

Classifier	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)	FPR (%)
J48	97.2818	97.4	97.3	97.1	1.8
Bayes Net	99.5708	99.6	99.6	99.6	1
Naive Bayes	99.05	100	100	100	0.03

After applying the 16 relevant features, the results obtained before the feature selection were not changed after the feature selection. J48 Decision Tree demonstrated an accuracy detection rate of 97.2818%, a precision of 97.4%, a recall of 97.3%, an F-measure (F1) of 97.1%, and a low false positive rate of 1.8% for all 33 features. Bayes Net demonstrated an accuracy detection rate of 99.5708%, a precision of 99.6 %, a recall of 99.6%, an F-measure (F1) of 99.6%, and a low false positive rate of 1.0% for all 33 features. Naive Bayes demonstrated an accuracy detection rate of 99.95 %, a precision of 100 %, a recall of 100 %, an F-measure (F1) of 100 %, as well as a low false positive rate of 0.03% for all 33 features shown in Tab. 4. These results indicate that all three algorithms effectively identify anomalous behaviors on the IoT network and differentiate regular traffic from attack traffic with high accuracy, precision, recall, and F-measure. The naive Bayes algorithm achieved the highest accuracy, which is 99.95%. Also, algorithms achieved gains in confidentiality, integrity, and availability when applied to the Redmi Note 9S device.

5 CONCLUSION

This research discussed machine learning and what solutions it can provide to industry, specifically IoT devices, to maintain information confidentiality, integrity, and availability while protecting users' privacy. Chosen machine learning algorithms J48, Bayes Net, and Naive Bayes were proposed solutions to investigate the unsolicited traffic and classify the traffic as normal and abnormal to safeguard IoT devices. The implementation was evaluated to see how effectively the three algorithms detect unsolicited traffic and classify the traffic. The Naive Bayes achieved a high accuracy detection rate of 99.95%, a precision of 100%, a recall of 100%, an F-measure (F1) of 100%, and a low false positive rate of 0.03%. Also, the Naive Bayes showed that it performs variable feature selection implicitly. To conclude, the Decision Tree (J48), Bayes Net, and Naive Bayes were practical algorithms for securing IoT devices and preventing privacy invasion. Naive Bayes is the best among the three used algorithms in IoT home networks.

6 REFERENCES

[1] Navani, D., Jain, S., & Nehra, M. S. (2017, December). The Internet of Things (IoT): A study of architectural elements. In *IEEE 13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS2017)*, 473-478. https://doi.org/10.1109/SITIS.2017.83

[2] Romkey, J. (2016). Toast of the IoT: the 1990 interop internet toaster. *IEEE Consumer Electronics Magazine*, *6*(1), 116-119. https://doi.org/10.1109/MCE.2016.2614740

[3] Ashton, K. (2009). That 'Internet of Things' thing. *RFID Journal*, *22*(7), 97-114.

[4] Tewari, A., & Gupta, B. B. (2020). Security, privacy, and trust of different layers in the Internet-of-Things (IoT) framework. *Future generation computer systems*, *108*, 909-920. https://doi.org/10.1016/j.future.2018.04.027

[5] Mohanta, B. K., Jena, D., Satapathy, U., & Patnaik, S. (2020). Survey on IoT security: Challenges and solution using machine learning, artificial intelligence and blockchain technology. *Internet of Things*, *11*, 100227. https://doi.org/10.1016/j.iot.2020.100227

[6] Mahmoudi, R., Roozi, S., Saghiri, A. M., & Mahmoudi, A. (2020). Extracting strategies for improving internet-of-things-based home industries in Iran: A strengths, weaknesses, opportunities, and threats analysis. *IEEE Transactions on Engineering Management*, *68*(2), 586-598. https://doi.org/10.1109/TEM.2020.2991859

[7] Varol, N., Aydogan, A. F., & Varol, A. (2017, April). Cyber attacks targeting Android cellphones. In *IEEE 5th International Symposium on Digital Forensic and Security (ISDFS2017)*, 1-5. https://doi.org/10.1109/ISDFS.2017.7916511

[8] Mohanta, B. K., Jena, D., Ramasubbareddy, S., Daneshmand, M., & Gandomi, A. H. (2020). Addressing security and privacy issues of IoT using blockchain technology. *IEEE Internet of Things Journal*, *8*(2), 881-888. https://doi.org/10.1109/JIOT.2020.3008906

[9] Carayannis, E. G., Grigoroudis, E., Rehman, S. S., & Samarakoon, N. (2019). Ambidextrous cybersecurity: The seven pillars (7Ps) of cyber resilience. *IEEE transactions on engineering management*, *68*(1), 223-234. https://doi.org/10.1109/TEM.2019.2909909

[10] Hussain, F., Hussain, R., Hassan, S. A., & Hossain, E. (2020). Machine learning in IoT security: Current solutions and future challenges. *IEEE Communications Surveys & Tutorials*, *22*(3), 1686-1721. https://doi.org/10.1109/COMST.2020.2986444

[11] Tournier, J., Lesueur, F., Le Mouël, F., Guyon, L., & Ben-Hassine, H. (2021). A survey of IoT protocols and their security issues through the lens of a generic IoT stack. *Internet of Things*, *16*, 100264. https://doi.org/10.1016/j.iot.2020.100264

[12] Talari, S., Shafie-Khah, M., Siano, P., Loia, V., Tommasetti, A., & Catalão, J. P. (2017). A review of smart cities based on the Internet of Things concept. *Energies*, *10*(4), 421. https://doi.org/10.3390/en10040421

[13] Garba, A. A., Siraj, M. M., Othman, S. H., & Musa, M. A. (2020). A study on cybersecurity awareness among students in Yobe State University, Nigeria: A quantitative approach. *Int. J. Emerg. Technol*, *11*(5), 41-49.

[14] Goyal, P., & Goyal, A. (2017, September). Comparative study of two most popular packet sniffing tools-Tcpdump and Wireshark. In *IEEE 9th International Conference on Computational Intelligence and Communication Networks (CICN2017)*, 77-81. https://doi.org/10.1109/CICN.2017.8319360

[15] Nam, S. Y., Djuraev, S., & Park, M. (2013). Collaborative approach to mitigating ARP poisoning-based Man-in-the-Middle attacks. *Computer Networks*, *57*(18), 3866-3884. https://doi.org/10.1016/j.comnet.2013.09.011

[16] Al Sukhni, B., Mohanta, B. K., Dehury, M. K., & Tripathy, A. K. (2023, July). A Novel Approach for Detecting and Preventing Security attacks using Machine Learning in IoT. In *IEEE 14th International Conference on Computing Communication and Networking Technologies (ICCCNT2023)*. 1-6. https://doi.org/10.1109/ICCCNT56998.2023.10307883

[17] Moustafa, N., Turnbull, B., & Choo, K. K. R. (2018). An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things. *IEEE Internet of Things Journal*, *6*(3), 4815-4830. https://doi.org/10.1109/JIOT.2018.2871719

[18] Al Sukhni, B., Dave, J. M., Manna, S. K., & Zhang, L. (2022, December). Investigating the security issues of multi-layer IoT attacks using machine learning techniques. In *IEEE Human-Centered Cognitive Systems (HCCS2022)*, 1-9. https://doi.org/10.1109/HCCS55241.2022.10090400

[19] Salman, O., Elhajj, I. H., Chehab, A., & Kayssi, A. (2022). A machine learning based framework for IoT device identification and abnormal traffic detection. *Transactions on Emerging Telecommunications Technologies*, *33*(3), e3743. https://doi.org/10.1002/ett.3743

[20] Cabrera, A., & Calix, R. A. (2016, October). On the anatomy of the dynamic behavior of polymorphic viruses. In *IEEE International Conference on Collaboration Technologies and Systems (CTS2016)*. 424-429. https://doi.org/10.1109/CTS.2016.0081

[21] Al Sukhni, B., Manna, S. K., Dave, J. M., & Zhang, L. (2022, December). Machine learning-based solutions for securing IoT systems against multilayer attacks. In *International Conference on Communication, Networks and Computing* (pp. 140-153). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-43140-1_13

[22] Lopez-Martin, M., Carro, B., Sanchez-Esguevillas, A., & Lloret, J. (2017). Network traffic classifier with convolutional and recurrent neural networks for Internet of Things. *IEEE Access*, *5*, 18042-18050. https://doi.org/10.1109/ACCESS.2017.2747560

Author's contacts:

Badeea Al Sukhni
Department of Cyber Security, University of Westminster,
309 Regent Street, W1B 2HW London, United Kingdom
E-mail: badeaa.alsukhni94@gmail.com

Bhabendu Kumar Mohanta
(Corresponding author)
College of Information Technology, United Arab Emirates University,
Al Ain, P.O. Box 15551, United Arab Emirates
E-mail: bhabendu@uaeu.ac.ae

Debnath Bhattacharyya
Department of Information Technology,
Aditya Institute of Technology and Management,
Tekkali, Andhra Pradesh, 532201, India
E-mail: debnathb@gmail.com

Tai-hoon Kim
(Corresponding author)
School of Electrical and Computer Engineering,
Yeosu Campus, Chonnam National University,
50, Daehak-ro, Yeosu-si, Jeollanam-do, 59626, Republic of Korea
E-mail: taihoonn@chonnam.ac.kr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240514171443

Original scientific paper
Received: 2024-05-14, Accepted: 2024-11-27

Empirical Assessment of Audience Attraction in Shopping Centres Based on the Notion of Social Sustainability

Reza Askarizad*, Roomina Soleymani Ardejani, Hossein Safari

Abstract: This study addresses the contemporary focus of commercial spaces on enhancing social, cultural, and economic experiences to attract individuals, prompting architects to align designs with audience objectives. It explores social sustainability dimensions, encompassing physical and social attributes, to create an appealing, human-centric shopping centre. Research involves gathering and analysing audience opinions through a distributed questionnaire, with SPSS software utilized for analysis. Findings reveal audience prioritization of social factors over physical ones, with interior decoration, spatial configuration, and proper lighting emerging as top physical preferences. In the social realm, satisfaction, social justice, and social security receive the most attention. Beyond theoretical implications, the study offers practical insights for designing sustainable shopping centres, emphasizing real-world applications based on social sustainability principles. The findings recommend a focus on creating vibrant, secure, and sociable atmospheres to attract a diverse audience, presenting a unique contribution to the field.

Keywords: shopping centre; social sustainability; spatial configuration; sustainable architecture; sustainable development

1 INTRODUCTION

As urban development progresses and society experiences an improvement in socio-cultural levels, small shopping centres no longer suffice to meet societal requirements. Consequently, shopping centres must re-evaluate their design processes, aligning them with the principles of social sustainability to cater to the contemporary needs of the populace. This necessitates a focused consideration of cultural, social, and economic aspects, ensuring these centres serve their functions effectively and capture community attention. Architects traditionally drew inspiration from existing markets and bazaars when designing shopping malls. However, what received less attention during this design phase were the social spaces within markets, encompassing a myriad of social, cultural, religious, and even political activities. Consequently, studies conducted by sociologists and psychologists revealed that visitors to shopping centres have motivations beyond mere shopping and they tend to be engaged in diverse activities [1].

Social sustainability, within the context of this study, is the facet of sustainable design that emphasizes the human dimension of the sustainability discourse. It focuses on creating schemes and environments that meet present-day needs without compromising resources for future generations, with a specific emphasis on the well-being of individuals and communities. Researchers have acknowledged the need to transform shopping malls from purely commercial spaces into multifunctional centres that accommodate both shopping and leisure activities [2]. Commercial centres play a crucial role in fostering the socio-cultural development of communities, contributing to their vibrancy and sociability. Given that shopping centres attract diverse audiences and serve as hubs for dynamic social interactions, integrating the principles of social sustainability during the planning phase enhances audience presence and continuity. Consequently, when designing shopping centres with a focus on social sustainability, careful attention is given to human behaviours and requirements, establishing a lasting relationship between individuals and their environment. Implementing appropriate procedures in the design and construction of shopping malls creates favourable conditions for the advancement of social, environmental, and economic considerations.

Commerce plays a crucial role in promoting the economic growth of a city, and the commercial space is recognized as one of the city's most vital components. Beyond its economic function, it serves as a hub for social interactions. In the current era, characterized by growing population trends and heightened consumer expectations, the proliferation of shopping centres throughout the city is seen as a fitting response to the audience's needs. Consequently, the demand for commercial spaces and shopping centres that not only address consumer demands but also provide for their social needs is more pronounced than ever, particularly in developing countries. However, the haphazard distribution of commercial centres in the city, coupled with undesirable designs and a lack of awareness regarding accessible routes, has resulted in a significant waste of time and energy in Iran. Consequently, to meet the evolving demands of societies, the interplay between factors influencing social growth and physical quality becomes increasingly evident. In this study, the focus is on prioritizing factors that contribute to an appealing public space, aiming to meet the needs and optimize people's leisure time. Specifically, the dimensions of social sustainability, encompassing both the physical and social environment, are examined as part of the study's particular scope. To guide the investigation, the following questions are proposed, seeking appropriate answers.

- In designing a shopping centre with a social sustainability approach, which of the social and physical factors will have the greatest impact on attracting the audience?
- How can the social and physical components of sustainability be intermingled and applied in the design of a shopping centre?

In this study, it is hypothesized that certain social factors, such as liveability and security, will have a significant impact on attracting visitors to business centres. Additionally, the spatial configuration structure of shopping malls, assumed as a moderating variable, can establish a constructive interaction between social and physical factors. By recognizing and adopting indicators of social sustainability, encompassing both physical and social aspects, a more attractive business environment will be provided to the audiences. Historically, the majority of studies related to shopping centres have concentrated on either socio cultural aspects [3-5] or physical factors [6-8] so far. While studies on the combination of social and physical factors affecting shopping centres simultaneously have not been sufficiently addressed. Thus, the present study aimed to bridge the identified scientific gap. The implications of this study can be fruitful for policymakers, architects, urban designers, and planners.

2 LITERATURE REVIEW

To realize the multifaceted nature of social sustainability, it is indispensable to define its dimensions within the realm of shopping centres. The concept encompasses both physical and social elements, creating a dynamic framework for evaluating the factors that contribute to an attractive and sustainable shopping environment.

2.1 Physical Factors Affecting Audience Attraction

Creating an appealing shopping centre goes beyond aesthetics and involves careful consideration of various design elements. On the one hand, interior design, a critical aspect, not only triggers positive emotional responses but also fosters a connection between consumers and the retail space [9]. Aesthetically pleasing interiors encourage visitors to explore their surroundings, spending more time in the shopping centre [10]. On the other hand, spatial configuration plays a crucial role in shaping the overall customer experience to meet their socio-functional requirements [11]. Well-organized spatial configurations enhance navigation, reduce frustration, and create a sense of ease for shoppers [12]. Moreover, accessibility and convenient traffic flow contribute significantly to a positive perception of the shopping centre's functionality [13]. The choice of colour schemes and textures also profoundly impacts consumer behaviour and evokes specific emotions and influences purchasing decisions [14]. For instance, the choice of warm colours like red and orange in retail spaces, known to stimulate excitement and sociability, can enhance the shopping experience [15]. In addition, incorporating cool tones such as blue and green can create a calm and inviting atmosphere, encouraging relaxed browsing and increasing the time spent in the store [16]. Previous studies have emphasized that texture adds a tactile dimension, making the shopping environment more engaging [17]. Furthermore, noise pollution should be minimized for a serene shopping environment. A soundproof structure reduces distractions, enhancing the overall acoustic experience, and correlates with increased customer satisfaction [18]. Additionally, optimal thermal comfort, maintained through proper heating, ventilation, and air conditioning systems, is integral to ensuring visitor satisfaction and building sustainability [19]. Considering the psychological impact of lighting is considered an indispensable factor in audience attraction of public spaces. On one hand, proper lighting not only serves a practical purpose but also enhances the mood and atmosphere within the shopping centre [20]. Strategic lighting design draws attention to specific areas, creating focal points that guide visitors and contribute to an aesthetically pleasing environment [21]. On the other hand, integrating greenery and natural elements aligns with the trend of environmentally conscious design [22]. Moreover, beyond visual appeal, green spaces contribute to a sustainable and pleasant environment [23] and positively influence visitors' psychological well-being [24]. Each of these elements contributes synergistically to create a shopping environment that not only appeals to the senses but also enhances the overall customer experience based on physical attributes (Fig. 1).

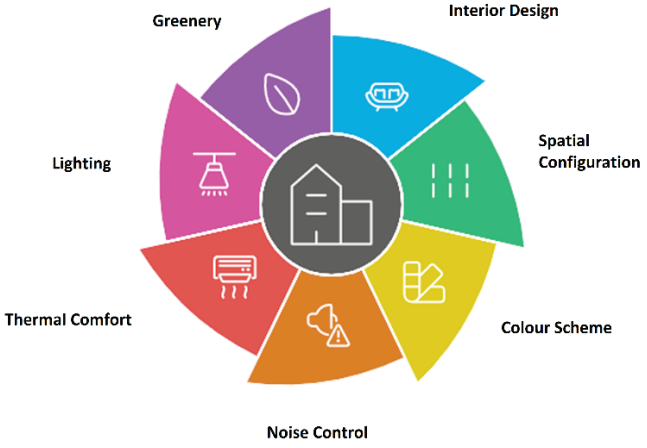


Figure 1 Influential physical factors on enhancing shopping centre appeals based on conducted literature [9-24]

2.2 Social Factors Affecting Audience Attraction

At the core of audience attraction in public spaces lies the concept of satisfaction. Research consistently highlights the link between overall satisfaction and the likelihood of repeat visits. Understanding and addressing the factors that contribute to customer satisfaction are paramount for the sustained success of a shopping centre [25]. In an era where consumers value socially responsible business practices, the concept of social justice becomes a key determinant of shopping centre attractiveness [26]. Infusing liveability into a shopping centre involves creating a dynamic and energetic atmosphere through initiatives such as cultural events, interactive activities, and entertainment programs. Successful strategies involve organizing events and activities that go beyond traditional retail, transforming the space into a lively and engaging environment [3]. Shopping centres are not merely transactional spaces; they are evolving into community hubs. Facilitating social interactions through

events and collaborative spaces contributes to a vibrant and lively atmosphere, attracting a diverse range of visitors [27]. Studies suggest that sustaining social interaction in urban environments extends beyond verbal communication, encompassing deeper emotional and cultural connections within individuals [28]. In addition, the perceived level of security within a shopping centre significantly influences consumer behaviour. Visible security measures create a sense of safety, fostering trust and positively impacting the overall shopping experience [5]. Empowering visitors through social participation initiatives, such as feedback mechanisms and community involvement, creates a sense of belonging and ownership [29]. This engagement contributes to the continuous improvement of the shopping centre based on the community's needs and preferences. Creating a sense of belonging through social participation initiatives fosters a deeper connection between visitors and the shopping centre, transforming it into a community-focused space [30]. The establishment of a community identity encourages shoppers to view the centre as more than just a commercial venue, but as a destination where they belong [31]. The articulated process of social factors influencing audience attraction in shopping centres is summarized in a graphical presentation, as shown in Fig. 2.

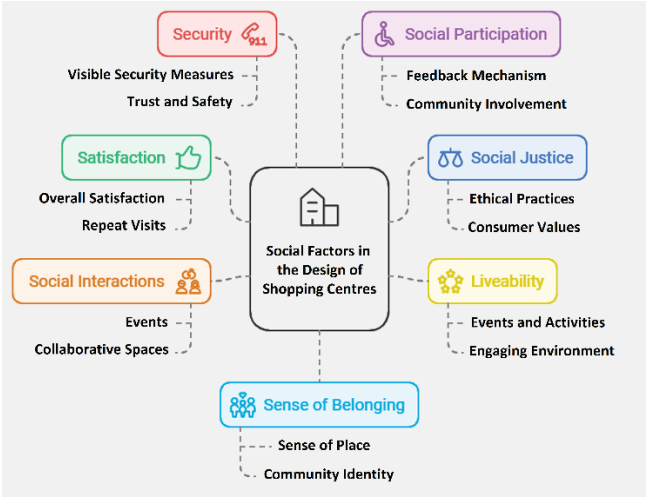


Figure 2 Graphical representation of the social factors affecting audience attraction in shopping centres based on conducted literature [25-31]

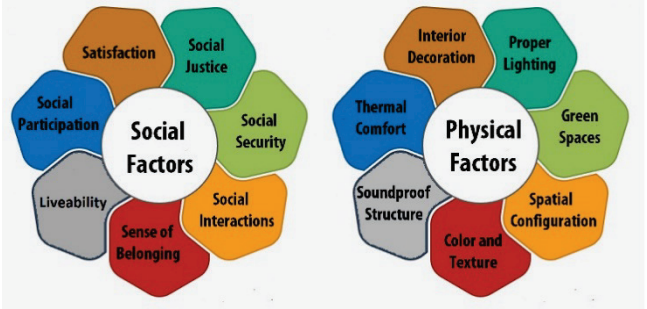


Figure 3 Factors influencing audience attractions towards shopping centres based on the notion of social sustainability based on the reviewed literature [9-31]

In sum, this reviewed literature outlines the theoretical framework for understanding the factors affecting audience attraction in shopping centres based on the notion of social sustainability. By exploring both the physical and social dimensions, the research seeks to contribute to the creation of shopping environments that not only address functional and emotional needs but also resonate with the values and preferences of the contemporary consumer. A summary of influencing factors is also depicted in the following diagram (Fig. 3).

3 MATERIAL AND METHODS

The study adopted an empirical research design to evaluate audience attraction in shopping centres, with a specific focus on social sustainability. Data was gathered through a structured field survey, conducted across multiple shopping centres selected for their diverse demographics and visitor frequency across northern Iran. The questionnaire, developed in a comparative and quantitative manner, served as the primary tool for collecting data on audience preferences and perceptions. A pre-test study involving 17 participants was conducted to refine and validate the questionnaire, ensuring its reliability using Cronbach's Alpha. For data collection, a structured questionnaire was employed. Utilizing the Cochran formula, the study aimed for a sample size of 384 respondents, featuring demographic representation with 252 females and 132 males. Participant diversity included varying age groups: 72% young, 15% middle-aged, 8% teenagers, and 4% elderly. Rigorous efforts were made to ensure a well-balanced sample in terms of gender, age, and shopping frequency.

The applied methods included a step-by-step approach to measure audience preferences. First, the questionnaire was designed based on prior literature, focusing on both physical and social factors. It should be noted that the quality of the designated questionnaire was verified through three professors holding PhD in architecture. Second, pilot testing was performed to enhance clarity and consistency. Third, data collection was carried out using a variety of methods including online through google forms and on-site during peak visitor hours to capture a broad representation of responses. Fourth, statistical methods such as descriptive and inferential analysis were applied to identify key patterns and correlations.

Data analysis was carried out using SPSS statistical software. The reliability of the questionnaire was assessed through various statistical tests, including the Friedman test, with a significance level set below 0.05. Examined variables included physical factors, encompassing interior decoration, spatial configuration, proper lighting, provision of green space, colour and texture, soundproof structure, and thermal comfort within shopping centres. Social factors were also explored, covering satisfaction, social justice, social security, social interaction, sense of belonging, liveability, and social participation. Ethical considerations were paramount in the study. Participants were fully informed about the research, and their informed consent was obtained. Measures were implemented to guarantee the confidentiality of participants' responses. It is important to note that the questionnaire surveys were developed following the guidelines of the

Declaration of Helsinki. As the survey ensured participant anonymity, voluntariness, and refrained from collecting data for participant identification, Institutional Review Board approval was not required for this study. To improve clarity, we have depicted the process of applied methods in this study in the form of a chart, presented as Fig. 4.

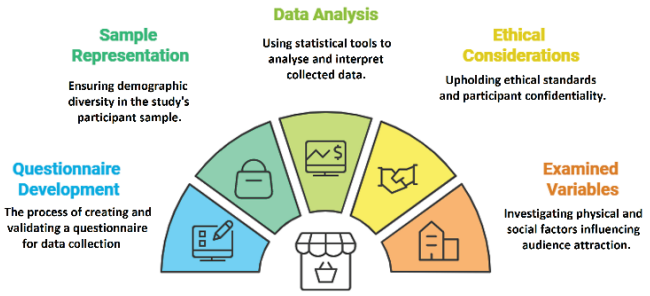


Figure 4 The step-by-step process of data collection and applied methods

4 RESULTS

In the present study, a questionnaire was utilized for data acquisition. The recorded demographic outcomes are reported as follows: 65% of the statistical population were women, and 34% were men. About 72% of the participants were young, 15% were middle-aged, 8% were teenagers, and 4% were elderly. The survey revealed significant patterns when analysing the correlation between demographic groups and usage duration. Young participants were the most frequent visitors, with 45% visiting once a week and 30% visiting once a month. In contrast, the middle-aged group primarily visited once a month (40%) or one to three times a year (25%). Similarly, teenagers showed a higher frequency of occasional visits, with 50% visiting once to three times a year. Elderly participants, representing a smaller portion of the statistical population, predominantly visited occasionally (40%) or once a month (30%). The results of the Friedman test, aimed at ranking the most influential priorities within physical factors affecting the design of shopping malls, indicated that interior decoration, with a mean rating of 5.11, was recognized as the most important physical factor. Subsequently, factors such as spatial configuration (4.61), proper lighting (4.28), and the provision of green space (3.79) were prioritized as the most important physical factors, respectively (Fig. 5). Findings obtained from the Friedman test also suggested that the most influential priority within social attributes was the satisfaction of the visitors, with a mean rating of 6.06. Following this, factors such as social justice (5.22), social security (4.69), and social interaction (4.25) were prioritized as the most important social attributes, respectively (Fig. 6).

In order to achieve the main objective established in this study concerning the most crucial priorities in terms of both physical and social attributes in the design of shopping centres, the Friedman test results revealed that satisfaction is a significant priority. Subsequently, factors such as social justice, interior decoration, social security, spatial configuration, proper lighting, social interaction, sense of belonging, liveability, and the provision of green space within shopping centres are considered the most important attributes based on the notion of social sustainability (Fig. 7).

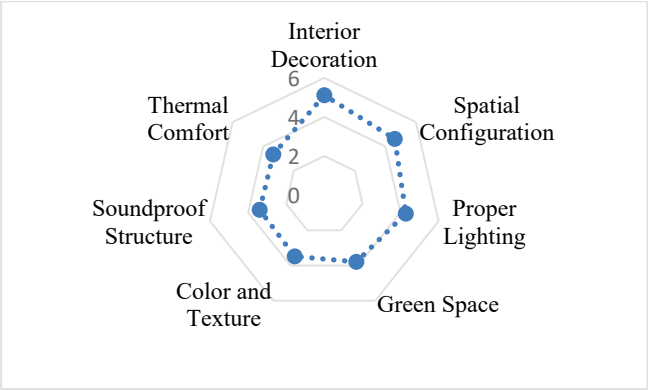


Figure 5 Friedman test results for prioritization of the physical attributes in the design of shopping centres

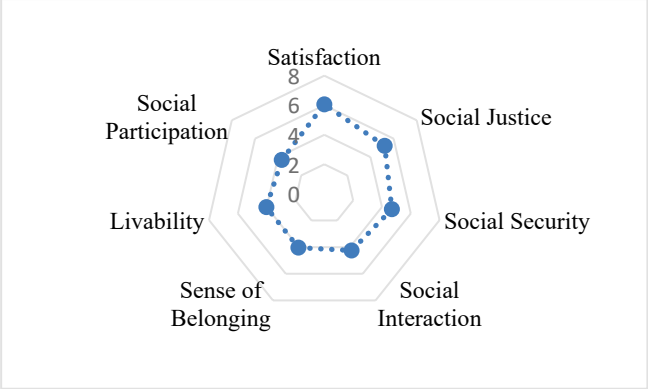


Figure 6 Friedman test results for prioritization of the social attributes in the design of shopping centres

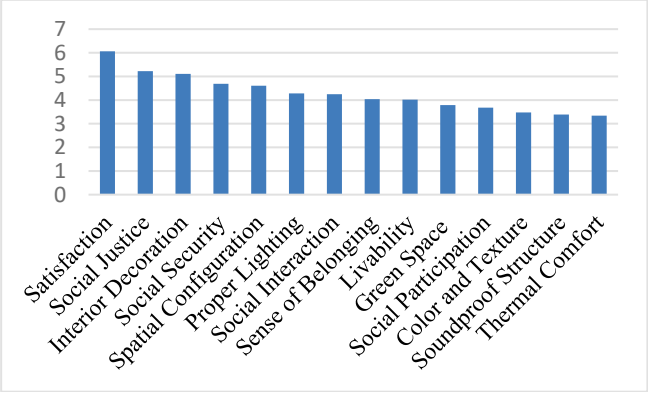


Figure 7 Friedman test results for prioritization of the social and physical attributes in the design of shopping centres

In summary, the findings have demonstrated that audiences prioritize social factors over physical ones. Demographic outcomes revealed notable trends, with a majority of the statistical population being women and a significant percentage representing the youth demographic. Findings related to usage duration underscored varying visit frequencies, with a substantial portion visiting monthly. The crucial aspect of visitor satisfaction emerged as a central theme, supported by a positive correlation between frequent visits and the prioritization of satisfaction. The Friedman test

ranked interior decoration as the foremost physical factor influencing shopping centre design, while social attributes were led by visitor satisfaction. The integration of both physical and social attributes as vital elements in the design, emphasizing social sustainability, emerged as a significant outcome.

5 DISCUSSION

The findings of this research are aligned with previously conducted studies with regard to the association between socialization and satisfaction as their primary priority [32]. In addition, previous studies have mainly focused on physical [10], social [33], and cultural [34] factors associated with the design of shopping centres. However, the current study adopts an incorporated approach to consider both physical and social factors that are interchangeably associated with each other to enhance the overall experience of the audiences within shopping centres. Consistent with previous studies that advocate the significant role of physical aspect of spatial configuration [35-37], this study broadened the vision towards incorporating social and physical attributes in architectural spaces. In contrast to preceding studies, our research uniquely contributes by emphasizing the practical implications of incorporating both social and physical factors into shopping centre design for sustainability. While previous studies predominantly rely on theoretical implications, our work provides valuable insights for designing sustainable shopping centres by emphasizing practical applications grounded in social sustainability principles. This novel approach enriches the existing literature, positioning our study as a noteworthy contribution to the field of shopping centre research. Accordingly, the following instructions for the pragmatic implications of shopping malls in terms of physical and social characteristics can be provided:

- Proper usage of natural and artificial lighting
- Utilization of collective and individual seating areas
- Appropriate spatial organization and configuration
- Adoption of resilient and flexible furniture arrangement
- Incorporation of vegetation, green space, and fountains
- Provision of platforms for attendance and interaction of people
- Playing relaxing and tranquil music
- Ensuring proper access to service spaces such as parking and restrooms
- Offering diverse products at reasonable prices
- Providing thermal and acoustic comfort in the interior space.

6 CONCLUSION

In this study, the main goal was to formulate a framework for enhancing audience attraction in shopping centers by incorporating social sustainability principles. Unlike many shopping malls in developing nations, which primarily cater to merchandise-related needs, there is a noticeable lack of infrastructure to address the broader social demands of contemporary shopping center audiences. Thus, the imperative to establish a framework emphasizing the interaction between social and physical attributes in shopping centers, grounded in the concept of social sustainability, becomes evident. Our findings highlight that key physical attributes in shopping centers encompass interior decoration, spatial configuration, and proper lighting, while critical social attributes include satisfaction, social justice, and social security, with social aspects taking precedence over physical ones according to the statistical population's perspective.

One restriction of the present study is its concentration in a developing country (Iran). Exploring cross-cultural contexts or other geographical areas could strengthen the research outcomes. Regarding suggestions for future research on the interaction of physical and social attributes in shopping centres, space syntax analysis can be implemented to assess social behaviours by scrutinizing spatial configuration characteristics. Further research could also delve into the evolving relationship between physical and online shopping, examining how digital technologies are reshaping the role of shopping centres and the experiences of shoppers. Overall, a conclusive framework has been delineated as guidance for the principles of audience attraction to shopping centres based on the notion of social sustainability (Fig. 8).

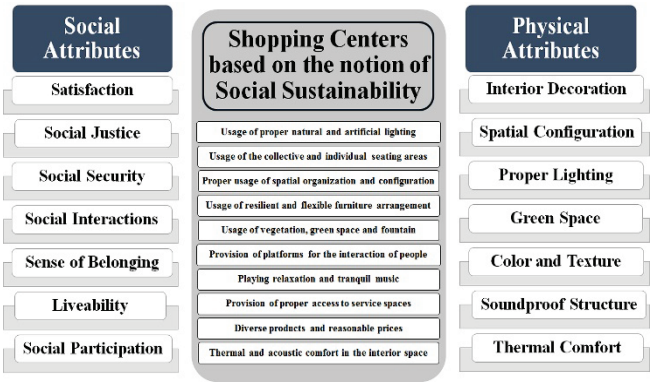


Figure 8 An instructional framework of the shopping centres based on the notion of social sustainability

In the context of shopping centres, the concept of social sustainability could be generalized by emphasizing the importance of meaningful and lasting connections among visitors. Beyond the transactional nature of shopping, creating an environment that fosters genuine social interactions and connections contributes to the sustained well-being of the community and the shopping centre itself. This could involve initiatives such as community events, social spaces within the centre, and engagement programs that go beyond the commercial aspect, promoting a sense of belonging and shared values among visitors. In conclusion, the main goal is to build a shopping centre atmosphere that goes beyond transactions and becomes a social hub, contributing to the overall social sustainability of the community it serves.

7 REFERENCES

[1] Wagner, T. (2007). Shopping motivation revised: a means-end chain analytical perspective. *International Journal of Retail & Distribution Management, 35*(7), 569-582. https://doi.org/10.1108/09590550710755949

[2] Iqbal, A., Nazir, H., & Memon, R. M. (2022). Shopping centers versus traditional open street bazaars: A comparative study of

user's preference in the city of Karachi, Pakistan. *Frontiers in Built Environment, 8*, 1066093. https://doi.org/10.3389/fbuil.2022.1066093

[3] Khalili, A., & NayyeriFallah, S. (2018). Role of social indicators on vitality parameter to enhance the quality of women's communal life within an urban public space (case: Isfahan's traditional bazaar, Iran). *Frontiers of Architectural Research, 7*(3), 440-454. https://doi.org/10.1016/j.foar.2018.06.001

[4] Carmelino, G., & Hanazato, T. (2019). The built environment of Japanese shopping streets as visual information on pedestrian vibrancy. *Frontiers of Architectural Research, 8*(2), 261-273. https://doi.org/10.1016/j.foar.2019.01.003

[5] Badiora, A. I., & Odufuwa, B.O. (2019). Fear dynamics in public places: a case study of urban shopping centers. *Journal of Place Management and Development, 12*(2), 248-270. https://doi.org/10.1108/JPMD-11-2018-0084

[6] Kusumowidagdo, A., Sachari, A., & Widodo, P. (2016). Visitors' perceptions on the important factors of atrium design in shopping centers: A study of Gandaria City Mall and Ciputra World in Indonesia. *Frontiers of Architectural Research, 5*(1), 52-62. https://doi.org/10.1016/j.foar.2015.11.003

[7] Ortegón-Cortázar, L., & Royo-Vela, M. (2017). Attraction factors of shopping centers: Effects of design and eco-natural environment on intention to visit. *European Journal of Management and Business Economics, 26*(2), 199-219. https://doi.org/10.1108/EJMBE-07-2017-012

[8] Fasli, M., Riza, M., & Erbilen, M. (2016). The Assessment and Impact of Shopping Centers: Case Study Lemar. *Open House International, 41*(4), 98-103. https://doi.org/10.1108/OHI-04-2016-B0013

[9] Burkut, E. B. (2023). Evaluation of interior architecture education research in the web of science database: Bibliometric and science mapping analysis. *Journal of Architectural Sciences and Applications, 8*(1), 385-405. https://doi.org/10.30785/mbud.1263207

[10] Arslan, H. D., & Ergener, H. (2023). Comparative analysis of shopping malls with different plans by using space syntax method. *Ain Shams Engineering Journal, 14*(9), 102063. https://doi.org/10.1016/j.asej.2022.102063

[11] Dogu, U., & Erkip, F. (2000). Spatial Factors Affecting Wayfinding and Orientation: A Case Study in a Shopping Mall. *Environment and Behavior, 32*(6), 731-755. https://doi.org/10.1177/00139160021972775

[12] Askarizad, R., & He, J. (2022). Perception of spatial legibility and its association with human mobility patterns: An empirical assessment of the historical districts in Rasht, Iran. International Journal of Environmental Research and Public Health, 19(22), 15258. https://doi.org/10.3390/ijerph192215258

[13] Zhang, L., Zhou, J., & Hui, E. C. M. (2020). Which types of shopping malls affect housing prices? From the perspective of spatial accessibility. *Habitat International, 96*, 102118. https://doi.org/10.1016/j.habitatint.2020.102118

[14] Jaglarz, A. (2023). Perception of Color in Architecture and Urban Space. *Buildings, 13*(8), 2000. https://doi.org/10.3390/buildings13082000

[15] Wasista, I. P. U., Kerdiati, N. L. K. R., & Darmastuti, P. A. (2022). The influence of color on consumer's affective side in commercial space: A literature review approach. *Journal of Aesthetics, Design, and Art Management, 2*(2). https://doi.org/10.58982/jadam.v2i2.263

[16] Elliot, A. J., Maier, M. A., Moller, A. C., Friedman, R., & Meinhardt, J. (2007). Color and psychological functioning: The effect of red on performance attainment. *Journal of Experimental Psychology: General, 136*(1), 154-168. https://doi.org/10.1037/0096-3445.136.1.154

[17] Faucheu, J., Weiland, B., Juganaru-Mathieu, M., Witt, A., & Cornuault, P.-H. (2019). Tactile aesthetics: Textures that we like or hate to touch. *Acta Psychologica, 201*, 102950. https://doi.org/10.1016/j.actpsy.2019.102950

[18] Soars, B. (2009). Driving sales through shoppers' sense of sound, sight, smell and touch. *International Journal of Retail & Distribution Management, 37*(3), 286-298. https://doi.org/10.1108/09590550910941535

[19] Dastoum, M., Guevara, C. S., & Arranz, B. (2024). Efficient daylighting and thermal performance through tessellation of geometric patterns in building façade: A systematic review. *Energy for Sustainable Development, 83*, 101563. https://doi.org/10.1016/j.esd.2024.101563

[20] Han, H., Sahito, N., Nguyen, T. V. T., Hwang, J., & Asif, M. (2019). Exploring the Features of Sustainable Urban Form and the Factors that Provoke Shoppers towards Shopping Malls. *Sustainability, 11*(17), 4798. https://doi.org/10.3390/su11174798

[21] Mayhoub, M. S., & Rabboh, E. H. (2022). Daylighting in shopping malls: Customer's perception, preference, and satisfaction. *Energy and Buildings, 255*, 111691. https://doi.org/10.1016/j.enbuild.2021.111691

[22] Safari, H., & Khaftani, P. S. (2022). The role of green playground on social interactions of children (Case study: Pasdaran Park in Bandar Anzali). *Tehnički glasnik, 16*(4), 464-470. https://doi.org/10.31803/tg-20220114174057

[23] Soleymani Ardejani, R., & Askarizad, R. (2018). Semiology of Semantic Components in Iranian Gardens Case Study: Eram and Narenjestan Gardens in Shiraz. *Creative City Design, 1*(2), 42-52. https://sanad.iau.ir/Journal/crcd/Article/919361

[24] Hong, S.-K., Lee, S.-W., Jo, H.-K., & Yoo, M. (2019). Impact of Frequency of Visits and Time Spent in Urban Green Space on Subjective Well-Being. *Sustainability, 11*(15), 4189. https://doi.org/10.3390/su11154189

[25] Amin, M., Ryu, K., Cobanoglu, C., Rezaei, S., & Wulan, M. M. (2021). Examining the Effect of Shopping Mall Attributes in Predicting Tourist Shopping Satisfaction and Behavioral Intentions: Variation across Generation X and Y. *Journal of Quality Assurance in Hospitality & Tourism, 22*(3), 367-394. https://doi.org/10.1080/1528008X.2020.1818667

[26] Newman, C., Rand, J., Tarp, F., & Trifkovic, N. (2020). Corporate Social Responsibility in a Competitive Business Environment. *The Journal of Development Studies, 56*(8), 1455-1472. https://doi.org/10.1080/00220388.2019.1694144

[27] Askarizad, R., Dastoum, M., & Garau, C. (2024). Street puppet theatre shows on the façades of commercial buildings as a novel stimulator for social gatherings in smart cities. *Buildings, 14*(9), 2950. https://doi.org/10.3390/buildings14092950

[28] Askarizad, R. (2020). Evaluation of the Effective Factors on Social Interactions in the Design of Public Libraries. *Tehnički glasnik, 14*(4), 403-410. https://doi.org/10.31803/tg-20190525220939

[29] Askarizad, R., He, J., & Soleymani Ardejani, R. (2022). Semiology of Art and Mysticism in Persian Architecture According to Rumi's Mystical Opinions (Case Study: Sheikh Lotf-Allah Mosque, Iran). *Religions, 13*(11), 1059. https://doi.org/10.3390/rel13111059

[30] Askarizad, R., Rezaei Liapee, S., & Mohajer, M. (2021). The Role of Sense of Belonging to the Architectural Symbolic Elements on Promoting Social Participation in Students within Educational Settings. *Space Ontology International Journal, 10*(4), 1-18. https://doi.org/10.22094/soij.2021.1930154.1419

[31] Kusumowidagdo, A., Sachari, A., & Widodo, P. (2015). Visitors' Perception towards Public Space in Shopping Center

in the Creation Sense of Place. *Procedia - Social and Behavioral Sciences, 184*, 266-272. https://doi.org/10.1016/j.sbspro.2015.05.090
[32] Kunc, J., Križan, F., Novotná, M., & Bilková, K. (2022). Social Dimension of Shopping Centers Operation: Managerial Perspectives. *Sustainability, 14*(2), 709. https://doi.org/10.3390/su14020709
[33] Hagberg, J., & Styhre, A. (2013). The production of social space: Shopping malls as relational and transductive spaces. *Journal of Engineering, Design and Technology, 11*(3), 354-374. https://doi.org/10.1108/JEDT-04-2011-0019
[34] Lin, M., Shih, S., & Perng, Y. (2020). Sustainable Shopping Mall Rehabilitation. *Sustainability, 12*(17), 6698. https://doi.org/10.3390/su12176698
[35] Khotbehsara, E. M., Safari, H., Askarizad, R., & Somasundaraswaran, K. (2022). Investigating the role of spatial configuration on visitors' spatial cognition in health-care spaces: case studies in Gilan, Iran. *Facilities, 40*(9/10), 617-637. https://doi.org/10.1108/F-07-2021-0067
[36] Soltani, S., Gu, N., Ochoa, J. J., & Sivam, A. (2022). The role of spatial configuration in moderating the relationship between social sustainability and urban density. *Cities, 121*, 103519. https://doi.org/10.1016/j.cities.2021.103519
[37] Sailer, K., & McCulloh, I. (2012). Social networks and spatial configuration: How office layouts drive social interaction. *Social Networks, 34*(1), 47-58. https://doi.org/10.1016/j.socnet.2011.05.005

Authors' contacts:

Reza Askarizad
(Corresponding author)
1) Department of Urban and Regional Planning, Universidad Politécnica de Madrid, 28040 Madrid, Spain
2) Department of Civil and Environmental Engineering and Architecture (DICAAR), University of Cagliari, Via Marengo 2, 09123 Cagliari, Italy
Reza.Askarizad@gmail.com
Reza.Askarizad@unica.it

Roomina Soleymani Ardejani
Academic Center for Education, Culture, and Research (ACECR),
Rasht 4196911383, Iran

Hossein Safari
Islamic Azad University of Rasht Branch,
Rasht 4196911383, Iran

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240627004731

Original scientific paper
Received: 2024-06-27, Accepted: 2024-11-22

Cybersecurity in Action: Unraveling the Effects of Individual, Social, and Organizational Determinants

Somayye Nasiri, Shahram Shahabi, Aynaz Shafiesabet, Maral Talebbeidokhti*, Elmira Alvandi Behineh

Abstract: This research aimed to develop and assess a framework for exploring how factors at the organizational, social, and individual levels influence cybersecurity behavior (CB), with a particular focus on the intermediary role played by awareness of cybersecurity. Participants in this investigation included 381 employees from various public sector entities, with data being gathered via a questionnaire. The analysis was conducted through SEM, specifically employing the PLS. Outputs from this study indicate that cybersecurity awareness (CA), security self-efficacy (SSE), the perceived threat severity, and the perceived vulnerability all positively and significantly impact CBs. Additionally, the research found that initiatives focused on security training and awareness, information security culture, social media influence, family and friends influence, and information systems knowledge all significantly contribute to enhancing CA. Similarly, these elements, alongside direct experience in information security, were shown to affec SSE, with the exception of the influence exerted by family and friends, which did not show a significant impact on it. Influences from social media, family and friends, and direct experiences with information security were also found to significantly affect the perceived severity of threats and perceived of vulnerability. Thus, this study concludes that the factors analyzed are strong indicators of behavior related to cybersecurity.

Keywords: cybersecurity awareness; cybersecurity behavior; information security culture; information systems knowledge; security self-efficacy

1 INTRODUCTION

As internet technology, mobile applications, and artificial intelligence continue to evolve in size and complexity, the nature of cyberattacks is becoming increasingly destructive, leading to unprecedented security challenges for society in cyberspace [1]. Ensuring adherence to security protocols poses a difficulty for institutions across both public and private sectors. Academics and industry experts acknowledge the importance of individuals in maintaining security; however, they are often identified as the most vulnerable point in the security framework due to their frequent failure to implement optimal security measures [2, 3].

Studies examining cybersecurity, including those conducted by Han et al. [4] and Ifinedo [5], have reported that employee behavior does not always align with established security policies. Certain employees may disregard the information security policies within their organization, while others may underestimate the risks associated with information security, even when provided with security policies and written guidelines. Notably, individuals who have received sufficient information security training within their companies do not consistently display heightened levels of cybersecurity behavior (CB), as highlighted by Ng and Xu [6]. Furthermore, a survey conducted by Cybersecurity Insiders in 2018 revealed that 90% of cybersecurity professionals believe their organizations are susceptible to insider threats. This data underscores the ongoing difficulty of achieving individual security compliance. Therefore, considering the importance of CBs of employees in reducing security threats to organizations, the present study tends to propose a model to examine the effect of organizational (ST&AP, information security culture (ISC)), social (social media influence, friends and family influence), and individual (information systems knowledge (ISK) and information security experiences (ISE)) factors on CB with the mediating role of cybersecurity awareness (CA), perceived threat severity (PTS), and perceived vulnerability (PV).

CA and CB: CA is crucial within an organization. CA determines organizational readiness for any cyber-attacks and the level of employee familiarity with the subject [7, 8]. Security awareness refers to the degree to which members of an organization comprehend the significance of information security, the level of security mandated by the organization, their personal security obligations, and their corresponding actions [9]. CA has been defined as a combination of awareness and taking specific actions to protect information system of the organization [10]. According to Hansche [11], emphasizing security awareness can heighten the significance of information systems security and the potential adverse impacts of a security breach or failure. Furthermore, cultivating security awareness is crucial for promoting desired individual security compliance behaviors. Prior research acknowledges the significance of CA, highlighting its potential to influence employee behavior positively. When employees possess a heightened awareness of cybersecurity, they are better equipped to comprehend cybersecurity threats and mitigate their adverse impact on the organization [8]. Hence, fostering a culture of CA within companies is vital to enhance employees' cybersecurity practices, thereby bolstering the organization's capacity to effectively combat cybersecurity threats. Siponen [12] contended that offering security awareness plays a pivotal role in influencing employees to alter their attitudes regarding cybersecurity. In a related vein, Donalds [13] discovered that CA directly correlates with cybersecurity compliance behavior. Research conducted by Simonet and Teufel [14], Li et al. [15], Donalds and Osei-Bryson [3], and Limna et al. [16] further supports the significance of CA in influencing CB. Thus, it can be inferred that:

H_1: CA is effective on employee CB.

Security Self-Efficacy (SSE) and CB: Self-efficacy has wide applications in different situations and is an important factor in human competence construction system. The

performance of individuals with comparable skills across diverse circumstances, ranging from low to high intensity, or of a single individual in varying conditions, is contingent upon their self-confidence [17]. Self-efficacy refers to beliefs and judgments people make about their capabilities to perform specific tasks in specific situations [18]. SSE refers to judgments people make about their abilities to perform cybersecurity countermeasures and is considered the capability to counter cybersecurity threats [6]. Self-efficacy plays a significant role in influencing various behaviors associated with information security, including the intention to adhere to information systems policies [19], cybersecurity-related behaviors [3, 8, 20], and ased on the research conducted by Johnston and Warkentin [21], it is evident that user intention to utilize anti-spyware software is influenced by certain factors. Hence, it can be inferred that:

H_2: SSE is effective on employee CB.

PTS, PV, and CB: Derived from social psychology, the Protection Motivation Theory stands out as a robust explanatory framework for forecasting users' inclination towards practicing security behaviors [22, 23]. Threat appraisal is a significant component of this theory. Information about threats plays a crucial role in risk recognition. Threat appraisal relates to user assessment of the risk level stemming from negligence in information security. This threat has the potential to compromise the integrity, confidentiality, and availability of information [15; 24]. PTS and PV are two important variables in threat appraisal.

The Protection Motivation Theory has been widely applied to explore employees' perceptions of cybersecurity threats and their development of coping strategies from diverse angles [25]. Threat appraisal involves the process through which individuals evaluate the level of risk associated with a menacing cybercrime. PV denotes the degree to which employees feel at risk from a cyber-attack and perceive a deficiency in preventive measures and tools [3, 8, 26]. Research conducted [15, 8, 24, 26] also indicates the role of PTS and PV in CB. As a result, we posit that:

H_3: PTS influences employee CB.

H_4: PV is effective on employee CB.

Security Training and CA Programs: Organizations allocate a significant portion of their budgets to educational and awareness programs in the field of IT security, as a lack of awareness and expertise in this area leads to security incidents [27]. There is a growing recognition that utilizing data-driven AI systems can enhance recruitment and retention efforts to diversify the workforce, similar to how organizations invest in educational programs to mitigate risks related to IT security due to a lack of awareness and expertise [28]. Consequently, organizations need security training and cybersecurity training and awareness programs, focusing on educating employees so they act according to certain IT security principles and objectives [27, 29]. Numerous organizations disseminate security policies and provide security training and awareness initiatives, which elucidate the appropriate usage and interaction with computers and internet-connected systems. These programs encompass various topics, including password management and phishing [14]. There are two distinct approaches to implementing such security measures. While some scholars propose employing a deterrent approach through fear-based campaigns, others advocate for skill-based actions, training, and awareness [14]. Research conducted by Simonet and Teufel [14], Burns et al. [30], and Wipawayangkool [9] underscores the significance of security training and awareness programs (ST&AP) in enhancing CA and self-efficacy. As a result, it is hypothesized that:

H_5: ST&AP are effective on CA.

H6: ST&AP are effective on SSE.

Information Security and Cybersecurity Culture: Promoting an ISC within organizations is a comprehensive approach to cybersecurity [14]. ISC positively impacts employee CB and ultimately enhances organizational potential to effectively counter cyber threats [8]. Schlienger and Teufel [31] argue that ISC should change employee values to boost internal drive for safe cyber behavior. Research findings [14, 32] also highlight the role of ST&AP in enhancing CA and SSE. Therefore, it is assumed that:

H_7: ISC is effective on CA.

H_8: ISC is effective on SSE.

Social Media, Friends, and Family Influence on Cybersecurity: One of the determinants related to CBs can be found in the social environment. Social influence emphasizes the way individuals view the significance of others' viewpoints [33]. The social influence exerted by peers, family, and media affects user intentions [34]. Particularly because the effects of cybersecurity incidents may not be immediately visible, anecdotes shared by friends and family can serve as alternative illustrations, enhancing the social learning process for cybersecurity concerns [14, 32]. Over the past few years, social media platforms have risen as the primary means for sharing and seeking information regarding protective behaviors. For example, Liu [35] discovered that seeking information via social media encouraged users to adopt more protective behaviors. Leadership styles, whether in project or operational environments [35], can significantly influence the social dynamics within organizations, including behaviors rooted in social influence and organizational citizenship, ultimately impacting economic growth and governance integrity [36]. Simonton and Teufel [14] also validated the influence of social media, friends, and family in the realm of cybersecurity. From this, we can deduce the following hypotheses:

H_9: The impact of social media is significant on CA.

H_{10}: The impact of social media is significant on SSE.

H_{11}: The impact of social media is significant on PTS.

H_{12}: The impact of social media is significant on PV.

H_{13}: Friends and family is effective on CA.

H_{14}: Friends and family is effective on SSE.

H_{15}: Friends and family is effective on PTS.

H_{16}: Friends and family is effective on PV.

ISK and Cybersecurity: Technical acumen signifies an organization's capacity or proficiency in utilizing scientific and technical knowledge via a series of procedures and techniques to enhance and innovate products and operations [37, 47]. Scholarly works have emphasized the importance of skills, abilities, and distinct competencies associated with technology, along with the acquisition and utilization of technological knowledge, because of the crucial impact they have on organizational effectiveness [38, 39, 40, 48]. Knowledge is closely linked to CA. Users with a good

understanding of cybersecurity can identify negative or positive situations when using technology. Simonton and Teufel [14] also confirmed the role of ISK in cybersecurity. Therefore, it is assumed that:

H_{17}: ISK is effective on CA.

H18: ISK is effective on SSE.

ISE and Cybersecurity: The security encounters and beliefs of employees, along with the training offered by their organizations, greatly influence the perceived seriousness of information security. When employees have prior experience in handling cybersecurity incidents, they are better equipped to thwart new cyberattacks [15]. Consequently, Employees who have undergone information security incidents will possess a greater degree of practical experience in addressing cybercrimes compared to those lacking such exposure. Consequently, their hands-on experience serves as a wellspring of knowledge, enabling them to more effectively implement cybersecurity measures [41]. Li et al. [15] also affirmed the influence of information security encounters on associated cybersecurity attitudes, including PTS, vulnerability, and SSE. As a result, it is postulated that:

H_{19}: ISE are effective on PTS.

H_{20}: ISE are effective on PV.

H_{21}: ISE are effective on SSE.

In summary, as highlighted, the significance of the variables discussed in this study regarding CBs of employees has been emphasized throughout the theoretical literature. However, empirical literature shows that few research efforts have been made to present a model for examining the effect of organizational, social, and individual factors on CB with the mediating role of CA. Hence, the current research mainly aims at introducing a model that analyzes the impact of organizational, social, and individual factors on cybersecurity practices, with CA acting as a mediator. Drawing from theoretical literature and a conceptual framework originating from existing research, Fig. 1 illustrates the proposed conceptual model. In this model, ST&AP, ISC, social media influence, friends and family influence, ISK, and ISE are considered as independent variables. CA, SSE, PTS, and PV are considered as mediating variables, and employee CB is considered as dependent variable.

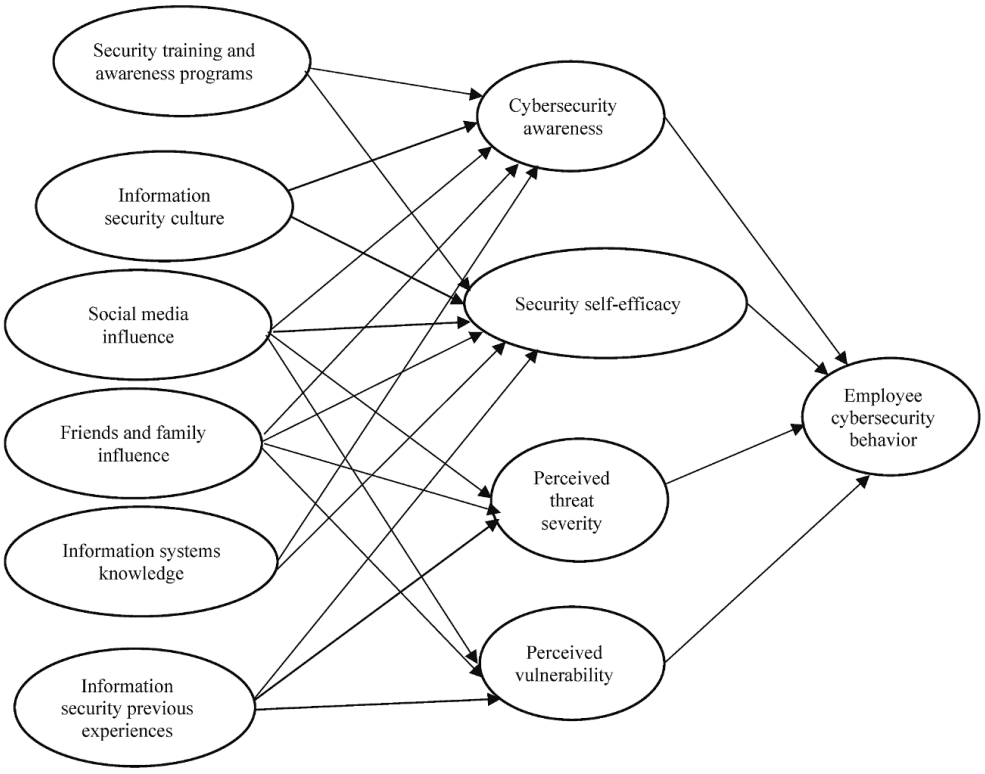


Figure 1 Conceptual model

2 RESEARCH METHODOLOGY

In the upcoming section, we will outline our research methodology.

2.1 Population and Sample

This study was conducted among Iranian public employees, among whom 450 questionnaires were distributed, with 409 returned and 28 removed due to incomplete answers, resulting in 381 questionnaires remained for analysis. Among the participants, 74% were male and 26% were female; 7% fell within the age range of 25-30 years, 41% between 31-40 years, 43% between 41-50 years, and 9% were aged 50 years or above.

2.2 Research Instruments

Questionnaires were used to measure the research constructs. The evaluation of ST&AP was conducted using the Simonet and Teufel [14] survey, which comprises two components. The evaluation of ISC was conducted utilizing

the Flores and Ekstedt [32] survey, which consists of four items. The assessment of social media influence was carried out using the Ng and Rahim [42] survey, which includes three items. The evaluation of friends and family influence was conducted using the Simonet and Teufel [14] survey, which comprises four items. To evaluate participants' knowledge of information systems, the Haeussinger & Kranz [43] questionnaire, comprising two items, was employed. The assessment of cyber security awareness utilized the Donalds and Osei-Bryson [3] survey, which includes three items. The assessment of previous experiences with information security was carried out using the Li et al. [15] survey instrument, which consists of three items measuring PTS, three items assessing PV, and three items evaluating SSE. The evaluation of CB was conducted using the Simonet and Teufel [14] survey, comprising nine items. Each item was assessed using a 5-point Likert scale.

3 RESULTS

3.1 Validity and Reliability of Instruments

To assess the validity and reliability of the research instrument, various tools were used, including Cronbach's alpha, composite reliability (CR), factor loadings, and the average variance extracted (AVE). CR utilizes factor loadings in its calculations, resulting in higher and improved values compared to those obtained from Cronbach's alpha [44]. The criterion for this index is also similar to Cronbach's alpha coefficient, which is 0.70 or higher for internal consistency of the measurement model. In confirmatory factor analysis, a factor loading of 0.60 or higher for any item signifies a clearly defined construct [45]. Tab. 1 shows that all factor loadings are above 0.6, all Cronbach's alpha and CR values are above 0.7, and all AVE values are above 0.5. According to Anbari et al. [46], an AVE value of 0.50 or higher suggests that the construct accounts for at least 50% of the variance in its associated markers.

Chin [45] proposes two criteria for assessing construct validity—each item should have a high factor loading on its own construct and a relatively low cross-loading on other constructs, with a suggested difference of at least 0.10 in factor loadings. Another criterion, the square root of the AVE should be greater than the correlations with other constructs. Based on the findings in Tab. 2, all variables exhibit the highest factor loading on their designated constructs, with a minimum difference of over 0.10 between their own factor loadings and those on other constructs, indicating satisfactory construct validity. Additionally, Tab. 3 presents results concerning correlations and the square root of the AVE.

Table 1 Results of measurement analysis

Variable	Item	Factor loading	Cronbach' alpha	CR	AVE
ST&AP	1	0.823	0.741	0.878	0.784
	2	0.944			
ISC	1	0.873	0.880	0.917	0.735
	2	0.863			
	3	0.857			
	4	0.837			
Social media influence	1	0.921	0.874	0.923	0.799
	2	0.889			
	3	0.872			
Friends and family influence	1	0.716	0.754	0.844	0.577
	2	0.715			
	3	0.828			
	4	0.773			
ISK	1	0.908	0.808	0.912	0.839
	2	0.924			
ISE	1	0.867	0.757	0.856	0.667
	2	0.856			
	3	0.718			
CA	1	0.860	0.782	0.874	0.699
	2	0.90			
	3	0.741			
SSE	1	0.891	0.858	0.913	0.779
	2	0.854			
	3	0.902			
PTS	1	0.853	0.924	0.952	0.870
	2	0.967			
	3	0.973			
PV	1	0.796	0.788	0.863	0.611
	2	0.781			
	3	0.809			
	4	0.739			
CB	1	0.820	0.929	0.941	0.638
	2	0.845			
	3	0.825			
	4	0.845			
	5	0.847			
	6	0.791			
	7	0.737			
	8	0.699			
	9	0.767			

Table 2 Cross factor loadings to assess validity of the questionnaires

	CA	CB	Family and Friends Influence	ISE	ISC	ISK	PTS	PV	ST&AP	SSE	Social Media Influence
CSA1	0.860	0.347	0.388	0.276	0.288	0.401	0.250	0.302	0.303	0.266	0.156
CSA2	0.900	0.476	0.454	0.243	0.372	0.442	0.301	0.330	0.286	0.393	0.154
CSA3	0.741	0.367	0.324	0.332	0.408	0.327	0.210	0.388	0.277	0.259	0.164
CSB1	0.362	0.820	0.468	0.380	0.547	0.448	0.521	0.492	0.441	0.402	0.414
CSB2	0.401	0.845	0.466	0.439	0.464	0.531	0.508	0.494	0.393	0.525	0.474
CSB3	0.387	0.825	0.428	0.396	0.469	0.457	0.504	0.515	0.427	0.443	0.449
CSB4	0.446	0.845	0.513	0.463	0.556	0.475	0.505	0.486	0.418	0.413	0.455
CSB5	0.446	0.847	0.533	0.431	0.575	0.500	0.577	0.558	0.414	0.502	0.375
CSB6	0.312	0.791	0.491	0.352	0.511	0.372	0.513	0.486	0.378	0.461	0.401

Table 3 Cross factor loadings to assess validity of the questionnaires (continuation)

	CA	CB	Family and Friends Influence	ISE	ISC	ISK	PTS	PV	ST&AP	SSE	Social Media Influence
CSB7	0.349	0.737	0.350	0.317	0.460	0.348	0.370	0.346	0.503	0.556	0.369
CSB8	0.353	0.699	0.324	0.342	0.430	0.315	0.381	0.350	0.437	0.498	0.379
CSB9	0.383	0.767	0.488	0.357	0.637	0.494	0.518	0.542	0.460	0.439	0.418
FFI1	0.336	0.370	0.716	0.397	0.379	0.361	0.343	0.484	0.308	0.301	0.314
FFI2	0.382	0.330	0.715	0.336	0.324	0.376	0.310	0.489	0.281	0.252	0.231
FFI3	0.316	0.526	0.828	0.475	0.543	0.494	0.531	0.540	0.439	0.373	0.464
FFI4	0.401	0.482	0.773	0.459	0.513	0.523	0.471	0.473	0.361	0.322	0.474
ISC1	0.337	0.592	0.459	0.467	0.873	0.530	0.525	0.531	0.436	0.458	0.448
ISC2	0.408	0.585	0.501	0.441	0.863	0.425	0.519	0.469	0.484	0.435	0.464
ISC3	0.343	0.540	0.500	0.465	0.857	0.482	0.540	0.539	0.464	0.375	0.509
ISC4	0.366	0.501	0.549	0.436	0.837	0.381	0.501	0.474	0.545	0.482	0.495
ISK1	0.409	0.452	0.528	0.393	0.414	0.908	0.503	0.500	0.288	0.413	0.367
ISK2	0.450	0.557	0.538	0.485	0.546	0.924	0.463	0.499	0.403	0.453	0.479
PE1	0.254	0.454	0.512	0.867	0.521	0.427	0.584	0.556	0.272	0.453	0.495
PE2	0.328	0.415	0.465	0.856	0.411	0.477	0.449	0.565	0.221	0.282	0.426
PE3	0.241	0.290	0.350	0.718	0.321	0.229	0.298	0.376	0.165	0.157	0.374
PS1	0.239	0.486	0.474	0.451	0.549	0.456	0.853	0.498	0.388	0.383	0.502
PS2	0.316	0.439	0.539	0.559	0.501	0.486	0.967	0.646	0.466	0.551	0.539
PS3	0.299	0.429	0.533	0.571	0.480	0.480	0.973	0.635	0.439	0.528	0.560
PV1	0.194	0.389	0.544	0.510	0.417	0.449	0.478	0.796	0.253	0.309	0.370
PV2	0.313	0.396	0.530	0.475	0.391	0.542	0.419	0.781	0.363	0.306	0.308
PV3	0.375	0.548	0.519	0.571	0.555	0.498	0.546	0.809	0.347	0.422	0.528
PV4	0.371	0.524	0.451	0.385	0.547	0.393	0.548	0.739	0.523	0.410	0.407
SET1	0.159	0.391	0.424	0.168	0.426	0.285	0.312	0.356	0.823	0.292	0.367
SET2	0.396	0.534	0.411	0.296	0.556	0.375	0.477	0.467	0.944	0.398	0.447
SMI1	0.102	0.457	0.423	0.531	0.490	0.398	0.454	0.492	0.428	0.291	0.921
SMI2	0.252	0.521	0.463	0.464	0.541	0.442	0.492	0.461	0.409	0.302	0.889
SMI3	0.152	0.414	0.450	0.434	0.466	0.408	0.487	0.451	0.408	0.262	0.872
SS1	0.332	0.516	0.365	0.319	0.477	0.433	0.447	0.414	0.396	0.891	0.283
SS2	0.285	0.486	0.362	0.390	0.401	0.363	0.452	0.405	0.268	0.854	0.274
SS3	0.365	0.551	0.369	0.336	0.476	0.451	0.497	0.417	0.386	0.902	0.288

Table 3 The results of construct validity

	CA	CB	Family and Friends Influence	ISE	ISC	ISK	PTS	PV	ST&AP	SSE	Social Media Influence
CA	0.836										
CB	0.480	0.799									
Family and Friends Influence	0.469	0.569	0.759								
ISE	0.334	0.486	0.553	0.817							
ISC	0.425	0.478	0.587	0.527	0.857						
ISK	0.470	0.553	0.582	0.482	0.528	0.916					
PTS	0.307	0.631	0.553	0.568	0.693	0.583	0.933				
PV	0.404	0.599	0.653	0.625	0.616	0.602	0.640	0.782			
ST&AP	0.344	0.535	0.462	0.277	0.565	0.380	0.463	0.474	0.885		
SSE	0.373	0.588	0.414	0.392	0.513	0.473	0.528	0.467	0.400	0.883	
Social Media Influence	0.188	0.520	0.497	0.534	0.558	0.465	0.572	0.524	0.464	0.319	0.894

3.2 SEM Results

Utilizing structural equation modeling (SEM), the study employed a proposed conceptual model and hypotheses to forecast the CB of public organization employees. The model was estimated using partial least squares (PLS), with t-values calculated through bootstrapping (500 subsamples). Fig. 2 visually depicts results of structural model. Tab. 4 reports the estimated path coefficients and Variance explained values.

According to the findings in Tab. 4, CA, SSE, PTS, and PV have a positive and significant impact on CB. Moreover, factors such as security training, awareness programs, ISC, social media influence, family and friends influence, and ISK demonstrate a positive and significant influence on CA.

Similarly, security training, awareness programs, ISC, social media influence, ISK, and ISE positively and significantly affect SSE. However, the influence of family and friends is not significant on SSE. Social media influence, family and friends influence, as well as ISE, have a positive and significant impact on PTS and vulnerability. The variables in the model explain 56% of the variability in CB, 33% in CA, 34% in SSE, 47% in PTS, and 54% in PV. Tab. 5 presents the indirect coefficients.

Tab. 5 indicates that CA mediates significantly the relationship between security training, awareness programs, ISC, social media influence, and family and friends influence on CB. Additionally, SSE acts as a significant and positive mediator between CA and security training, awareness programs, ISC, and social media influence in influencing CB. Moreover, PTS and vulnerability serve as significant and positive mediators between social media influence, family and friends influence, and ISE in impacting employee CB. However, the mediating role of SSE is not significant in the relationship between family and friends influence and CB. Tab. 6 presents the outcomes of hypothesis testing. The study found that the absolute Goodness of Fit (GOF) index for model is 0.571, indicating it is a well-suited fit.

4 DISCUSSION

This research aimed to introduce a model to investigate the impact of organizational, social, and individual factors on CB, incorporating the mediating role of CA through SEM. The findings indicated that the proposed model aligns well with the data collected for this study and can account for significant proportions of variance: 56% in CB, 33% in CA, 34% in cybersecurity self-efficacy, 47% in PTS, and 54% in PV.

The findings further revealed that ST&AP exert a positive and significant influence on CA and self-efficacy. Hence, these programs contribute to heightened CA and improved SSE. This outcome aligns with the research of Simonet and Teufel [14], Burns et al. [30], and Wipawayangkool [9]. It implies that when organizations offer training to enhance employees' understanding and awareness of computer and information security matters and educate them about their responsibilities in computer security, it results in increased CA and enhanced SSE. As a result, organizations should implement IT security-training programs aimed at augmenting employee knowledge and awareness of information technology security.

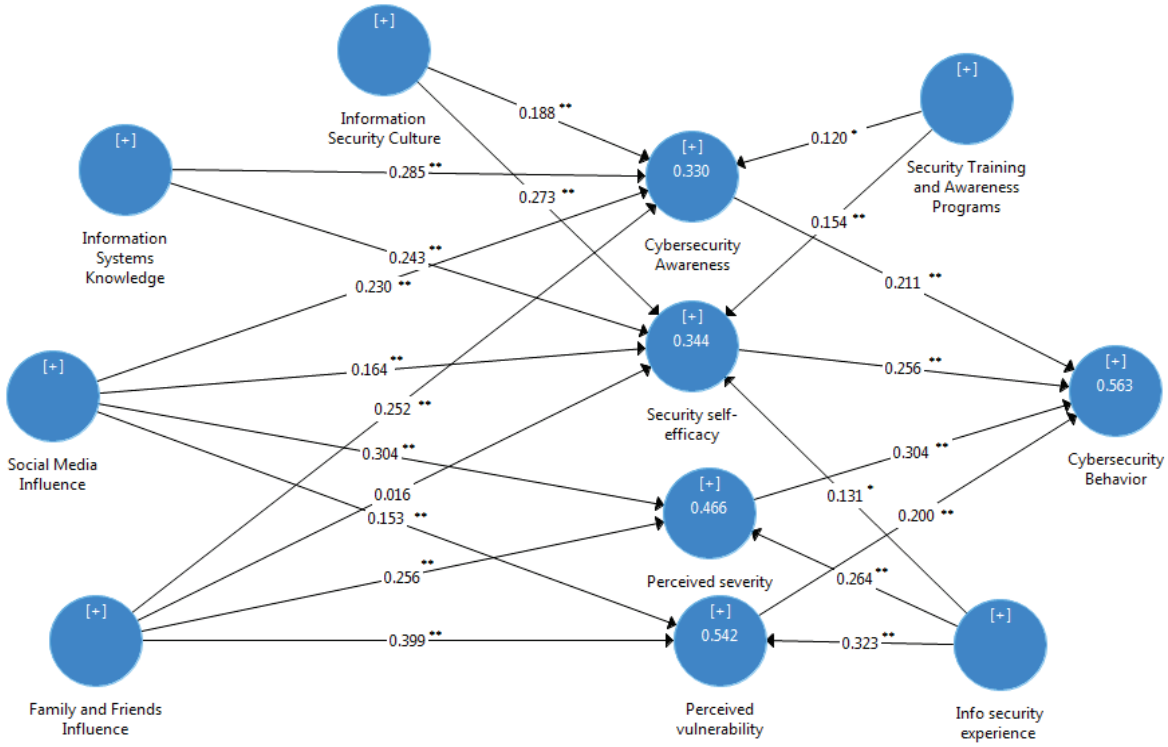


Figure 2 The tested model

Another finding of this study is the positive and significant effect of ISC on CA and SSE. Therefore, ISC leads to increased CA and SSE. This finding is consistent with Simonet and Teufel [14] and Flores and Ekstedt [32]. This finding can be attributed to the collective responsibility shared by employees within an organization to raise awareness about cybersecurity threats. When employees actively inform and caution one another about such threats, a common goal is established to protect information assets from unauthorized access and malware infection. This shared commitment leads to heightened CA and increased SSE among employees.

Furthermore, social media influences significantly CA, SSE, PTS, and PV. Therefore, social media influence leads to increased CA, SSE, PTS, and PV. This observation aligns with the work of Simonet and Teufel [14]. To clarify this result, if mass media encourages employees to practice safe cyber behaviors and reports influence employees to engage

in safe cyber behavior, employees feel pressured by mass media to practice safe cyber behavior, this will result in increased CA, SSE, PTS, and PV.

Table 4 Path coefficients and variance explained

Variable	β	t-value	p-value	Variance explained
On cybersecurity of employees via:				
CA	0.211**	4.818	0.01	0.563
SSE	0.256**	5.034	0.01	
PTS	0.304**	5.969	0.01	
PV	0.20**	4.270	0.01	
On CA via:				
ST&AP	0.12*	2.225	0.05	0.33
ISC	0.188**	2.820	0.01	
Social media influence	0.230**	3.989	0.01	
Friends and family influence	0.252**	4.034	0.01	
ISK	0.285**	4.249	0.01	
On SSE via:				
ST&AP	0.154**	3.036	0.01	0.344
ISC	0.273**	4.515	0.01	
Social media influence	0.164**	3.560	0.01	
Friends and family influence	0.016	0.223	0.01	
ISK	0.243**	5.149	0.01	
ISE	0.131*	2.382	0.01	
On PTS via:				
Social media influence	0.304**	6.056	0.01	0.466
Friends and family influence	0.256**	5.244	0.01	
ISE	0.264**	5.340	0.01	
On PV via:				
Social media influence	0.153**	3.058	0.01	0.542
Friends and family influence	0.399**	8.735	0.01	
ISE	0.323**	7.526	0.01	

*$p < 0.05$; **$p < 0.01$

Table 5 Indirect coefficients

Indirect paths	Indirect effects	t-value	p-values
Family and Friends Influence→CA→CB	0.053	3.070	0.002
ISC→CA→CB	0.040	2.268	0.024
ISK→CA→CB	0.060	3.312	0.001
ST&AP→CA→CB	0.025	1.987	0.047
Social Media Influence→CA→CB	0.049	3.101	0.002
Family and Friends Influence→PTS→CB	0.078	3.963	0.000
ISE→PTS→CB	0.080	4.279	0.000
Social Media Influence→PTS→CB	0.092	3.759	0.000
Family and Friends Influence→PV→CB	0.080	3.886	0.000
ISE→PV→CB	0.065	3.777	0.000
Social Media Influence→PV→CB	0.031	2.260	0.024
Family and Friends Influence→SSE→CB	0.004	0.214	0.831
ISE→SSE→CB	0.034	2.271	0.024
ISC→SSE→CB	0.070	3.203	0.001
ISK→SSE→CB	0.062	3.586	0.000
ST&AP→SSE→CB	0.039	2.493	0.013
Social Media Influence→SSE→CB	0.042	2.907	0.002

The study also found that the influence of friends and family positively and significantly affects CA, PTS, and PV. Therefore, their influence contributes to an increase in CA, PTS, and PV. This finding is consistent with the research by Simonet and Teufel [14]. It implies that when family members and friends encourage employees to practice safe cyber behavior and hold them accountable for doing so, it results in heightened CA, PTS, and PV. Since the consequences of cybersecurity incidents are not always observable or tangible, sharing stories about such incidents can reinforce the social learning process for cybersecurity issues and serve as alternative examples [14].

An additional outcome from the model indicated that possessing knowledge of information systems has a positive and significant impact on both CA and SSE. Therefore, ISK leads to an increase in CA and SSE of employees. This finding is consistent with Simonet and Teufel [14]. It suggests that general knowledge of employees about computers and the internet (e.g., web and email systems) can lead to an increase in their CA and SSE. In fact, employees with good knowledge of cybersecurity can identify negative or positive situations when using technology and their security behaviors can improve.

Table 6 Outcomes of hypothesis evaluation

Hypothesis	Result
H1: CA influences CB.	Confirmed
H2: SSE influences CB,	Confirmed
H3: PTS influences CB.	Confirmed
H4: PV influences CB.	Confirmed
H5: ST&AP influences CA.	Confirmed
H6: ST&AP influences SSE.	Confirmed
H7: ISC influences CA.	Confirmed
H8: ISC influences SSE.	Confirmed
H9: social media influences CA.	Confirmed
H10: social media influences SSE.	Confirmed
H11: social media influences PTS.	Confirmed
H12: social media influences PV.	Confirmed
H13: friends and family influences CA.	Confirmed
H14: friends and family influences SSE.	Rejected
H15: friends and family influence is effective on PTS.	Confirmed
H16: friends and family influence is effective on PV.	Confirmed
H17: ISK is effective on CA.	Confirmed
H18: ISK is effective on SSE.	Confirmed
H19: ISE is effective on PTS.	Confirmed
H20: ISE is effective on PV.	Confirmed

The findings also indicated that ISE have a positive and significant influence on SSE, PTS, and PV. This is consistent with the study by Li et al. [15], who also found that ISE impact PTS, PV, and SSE. Consequently, ISE contribute to an increase in SSE, PTS, and PV. This finding suggests that when employees undergo formal training on common computer security practices, their organization implements a clear information security policy, and security awareness training is provided to employees, it leads to heightened SSE, PTS, and PV.

Another significant finding of the research indicated a notable and positive correlation between CA and CB. Therefore, CA leads to an increase in CB. This finding is consistent with Simonet and Teufel [14], Li et al. [15], Donalds and Osei-Bryson [3], and Limna et al. [16]. Therefore, employee awareness of cybersecurity and its associated risks plays a determining role in their CB. Siponen [12] emphasizes that offering security awareness is crucial in motivating employees to modify their behaviors for enhanced cybersecurity measures.

The findings indicated a positive and significant correlation between SSE and CB. Therefore, SSE leads to an increase in CB. This finding is consistent with D'Arcy and Lowry [20], Donalds and Osei-Bryson [3]. This finding suggests that when employees feel assured in adjusting web

browser security settings, handling virus-infected files, and effectively removing spyware and malware from their computers, it leads to an improvement in CB.

Furthermore, the study revealed a notable and positive correlation between PTS and CB. In other words, when individuals perceive the severity of cybersecurity threats to be high, it results in an increase in their engagement in CB. This is consistent with the study by Li et al. [15] and Gerdenitsch et al. [24]. To explain this finding, if employees experience a virus infection on their computers due to opening suspicious email attachments or unauthorized access to their personal confidential information at the workplace, it becomes a serious concern for them. Additionally, the potential loss of information through hacking poses another significant problem. These factors contribute to an elevation in CB.

Another key finding from this research indicates that PV has a positive and significant impact on CB. In essence, a sense of vulnerability contributes to an increase in engaging in cybersecurity practices. This is consistent with the study by Li et al. [15] and Gerdenitsch et al. [24]. The study suggests that when employees perceive that adhering to their organization's information security policies reduces their vulnerability to security breaches, and recognize that non-compliance could make them susceptible to harmful attacks, they are more inclined to believe that their efforts safeguard organizational information against unauthorized access. Conversely, neglecting information security protocols puts organizational data and resources at risk, prompting individuals to be more proactive in practicing cybersecurity measures.

5 MANAGERIAL IMPLICATIONS

- Organizations should provide ST&AP to enhance employees' understanding and awareness of cybersecurity issues. This includes educating them on their responsibilities regarding computer security to improve their beliefs and behaviors in this area.
- Encouraging the establishment of an ISC within the organization can boost CA and SSE. Employees should be encouraged to share cybersecurity threat alerts, fostering a collective responsibility towards information security. A unified approach among employees is vital for protecting organizational assets against cyber threats.
- Offering planned ISE can empower employees to prevent new cyberattacks. Employees with prior ISE are better prepared to tackle cybercrimes compared to those without such exposure.
- Training employees in relevant cybersecurity practices and enhancing their competency in this domain can improve SSE, influencing various information security-related behaviors. Strengthening self-efficacy is crucial as it impacts different behaviors related to information security and cybersecurity.
- Accurate assessments of PTS and vulnerability are essential for shaping employee CBs. Understanding and evaluating cybersecurity risk levels accurately is critical to avoid any negligence in addressing security concerns effectively within the organization.

6 CONCLUSION

In conclusion, the variables examined in this study are strong indicators of employee CBs. CA plays a crucial mediating role, significantly impacting the influence of security training programs, ISC, social media, and personal networks on employee CB. Likewise, SSE mediates positively the impact of CA on security training, ISC, and social media on employee CB. Moreover, PTS and vulnerability also play important mediating roles, significantly affecting the influence of social media, personal networks, and ISE on employee CB. Consequently, considering these model variables is crucial for improving employee acceptance of CBs.

7 LIMITATIONS

This research focused solely on employees from public organizations in Iran, therefore generalizations may be limited. Additionally, the nature of the study is correlational, and causal inferences cannot be drawn from the relationships between model variables. The role of culture in findings of this study should not be overlooked, and further studies should explore the relationships between variables in different cultural contexts.

8 REFERENCES

[1] Darvishinia, N. (2023). AI in Education: Cracking the Code through Challenges: A Content Analysis of One of the Recent Issues of Educational Technology and Society (ET&S) Journal. *Partners Universal International Innovation Journal, 1*(4), 61-71. https://doi.org/10.5281/zenodo.8264262

[2] Samadi, H., Nazari-Shirkouhi, S., & Keramati, A. (2014). Identifying and analyzing risks and responses for risk management in information technology outsourcing projects under fuzzy environment. *International Journal of Information Technology & Decision Making, 13*(06), 1283-1323. https://doi.org/10.1142/S021962201450076X

[3] Keramati, A., Samadi, H., & Nazari-Shirkouhi, S. (2013). Managing risk in information technology outsourcing: an approach for analysing and prioritising using fuzzy analytical network process. *International Journal of Business Information Systems, 12*(2), 210-242. https://doi.org/10.1504/IJBIS.2013.052052

[4] Han, J., Kim, Y. J., & Kim, H. (2017). An integrative model of information security policy compliance with psychological contract: Examining a bilateral perspective. *Computers & Security*, *66*, 52-65. https://doi.org/10.1016/j.cose.2016.12.016

[5] Ifinedo, P. (2012). Understanding information systems security policy compliance: An integration of the theory of planned behavior and the protection motivation theory. *Computers & Security*, *31*(1), 83-95. https://doi.org/10.1016/j.cose.2011.10.007

[6] Ng, B. Y., & Xu, Y. (2007). Studying users' computer security behavior using the Health Belief Model. *PACIS 2007 Proceedings*, 45. https://aisel.aisnet.org/pacis2007/45

[7] Nazari-Shirkouhi, S., Miri-Nargesi, S., & Ansarinejad, A. (2017). A fuzzy decision making methodology based on fuzzy AHP and fuzzy TOPSIS with a case study for information

systems outsourcing decisions. *Journal of Intelligent & Fuzzy Systems, 32*(6), 3921-3943. https://doi.org/10.3233/JIFS-12495

[8] Ramezani, A. (2024, May). Fusion Models for Cyber Threat Defense: Integrating Clustering with Kmeans, Random Forests, and SVM against Windows Malware. In *IEEE World AI IoT Congress (AIIoT2024)*, 465-470) https://doi.org/10.1109/AIIoT61789.2024.10578947

[9] Wipawayangkool, K. (2009). Security awareness and security training: An attitudinal perspective. *SWDSI 2009*, 266-273.

[10] Kemper, G. (2019). Improving employees' cyber security awareness. *Computer Fraud & Security*, *2019*(8), 11-14. https://doi.org/10.1016/S1361-3723(19)30085-5

[11] Hansche, S. (2001). Designing a security awareness program: Part 1. *Information Systems Security, 9*(6), 14-22. https://doi.org/10.1201/1086/43298.9.6.20010102/30985.4

[12] Siponen, M. T. (2000). A conceptual foundation for organizational information security awareness. *Information management & computer security*, *8*(1), 31-41. https://doi.org/10.1108/09685220010371394

[13] Donalds, C. (2015). *Cybersecurity policy compliance: An empirical study of Jamaican government agencies.* Paper Presented at the SIG GlobDev Pre-ECIS Workshop.

[14] Simonet, J., & Teufel, S. (2019). The influence of organizational, social and personal factors on cybersecurity awareness and behavior of home computer users. In *ICT Systems Security and Privacy Protection: 34th IFIP TC 11 International Conference, SEC 2019, Lisbon, Portugal, June 25-27, 2019, Proceedings 34*, 194-208. Springer International Publishing. https://doi.org/10.1007/978-3-030-22312-0_14

[15] Li, L., He, W., Xu, L., Ash, I., Anwar, M., & Yuan, X. (2019). Investigating the impact of cybersecurity policy awareness on employees' cybersecurity behavior. *International Journal of Information Management*, *45*, 13-24. https://doi.org/10.1016/j.ijinfomgt.2018.10.017

[16] Limna, P., Kraiwanit, T., Siripipattanakul, S., Limna, P., Kraiwanit, T., & Siripipattanakul, S. (2023). The relationship between cyber security knowledge, awareness and behavioural choice protection among mobile banking users in Thailand. *International Journal of Computing Sciences Research*, *7*, 1133-1151. https://doi.org/10.25147/ijcsr.2017.001.1.123

[17] Schwarzer, R. (2014). *Self-efficacy: Thought control of action*. Taylor & Francis. https://doi.org/10.4324/9781315800820

[18] Bandura, A. (2006). Guide for constructing self-efficacy scales. *Self-efficacy beliefs of adolescents*, *5*(307-337).

[19] Bulgurcu, B., Cavusoglu, H., & Benbasat, I. (2010). Information security policy compliance: an empirical study of rationality-based beliefs and information security awareness. *MIS quarterly*, 523-548. https://doi.org/10.2307/25750690

[20] D'Arcy, J., & Lowry, P. B. (2017). Cognitive-affective drivers of employees' daily compliance with information security policies: A multilevel, longitudinal study. *Information Systems Journal*. https://doi.org/10.1111/isj.12173

[21] Johnston, A. C., & Warkentin, M. (2010). Fear appeals and information security behaviors: An empirical study. *MIS quarterly*, 549-566. https://doi.org/10.2307/25750691

[22] Aurigemma, S., & Mattson, T. (2018). Exploring the effect of uncertainty avoidance on taking voluntary protective security actions. *Computers & Security*, *73*, 219-234. https://doi.org/10.1016/j.cose.2017.11.001

[23] Chehrehpak, M., Alizadeh, A., & Nazari-Shirkouhi, S. (2018). An empirical study on factors influencing technology transfer using structural equation modelling. International *Journal of Productivity and Quality Management, 23*(3), 273-288. https://doi.org/10.1504/IJPQM.2018.089801

[24] Gerdenitsch, C., Wurhofer, D., & Tscheligi, M. (2023). Working conditions and cybersecurity: Time pressure, autonomy and threat appraisal shaping employees' security behavior. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *17*(4). https://doi.org/10.5817/CP2023-4-7

[25] Vance, A., Siponen, M., & Pahnila, S. (2012). Motivating IS security compliance: Insights from habit and protection motivation theory. *Information & Management*, *49*(3-4), 190-198. https://doi.org/10.1016/j.im.2012.04.002

[26] Li, L., Xu, L., & He, W. (2022). The effects of antecedents and mediating factors on cybersecurity protection behavior. *Computers in Human Behavior Reports*, *5*, 100165. https://doi.org/10.1016/j.chbr.2021.100165

[27] Kirova, D., & Baumöl, U. (2018). Factors that affect the success of security education, training, and awareness programs: A literature review. *Journal of Information Technology Theory and Application (JITTA)*, *19*(4), 4. https://aisel.aisnet.org/jitta/vol19/iss4/4

[28] Sadeghi, S., & Niu, C. (2024). Augmenting Human Decision-Making in K-12 Education: The Role of Artificial Intelligence in Assisting the Recruitment and Retention of Teachers of Color for Enhanced Diversity and Inclusivity. *Leadership and Policy in Schools*, 1-21. https://doi.org/10.1080/15700763.2024.2358303

[29] Nazari-Shirkouhi, S., Ansarinejad, A., Miri-Nargesi, S. S., Dalfard, V. M., & Rezaie, K. (2011). Information systems outsourcing decisions under fuzzy group decision making approach. *International Journal of Information Technology & Decision Making, 10*(06), 989-1022. https://doi.org/10.1142/S0219622011004683

[30] Burns, A. J., Roberts, T. L., Posey, C., Bennett, R. J., & Courtney, J. F. (2015, January). Assessing the role of security education, training, and awareness on insiders' security-related behavior: An expectancy theory approach. In *48th IEEE Hawaii International Conference on System Sciences*, 3930-3940. https://doi.org/10.1109/HICSS.2015.471

[31] Schlienger, T., & Teufel, S. (2002). *Information security culture.* In: Ghonaimy, M. A., El-Hadidi, M. T., & Aslan, H. K. (eds.) *Security in the Information Society (IAICT2002), 86*, 191-201. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-35586-3_15

[32] Flores, W. R., & Ekstedt, M. (2016). Shaping intention to resist social engineering through transformational leadership, information security culture and awareness. *Computers & security*, *59*, 26-44. https://doi.org/10.1016/j.cose.2016.01.004

[33] Tehranian, K., Khorsand, M. S., Zarei, M., Arani, G. G., Banabari, H. G., & Sasani, F. (2024). Unveiling the Impact of Social Media Usage on Firm Performance: The Mediating Influence of Organizational Agility and Innovation Capability. *Tehnički glasnik, 18*(4), 540-548. https://doi.org/10.31803/tg-20230918233848

[34] Gao, L., & Bai, X., (2014), A unified perspective on the factors influencing consumer acceptance of internet of things technology. *Asia Pacific Journal of Marketing and Logistics, 26*(2), 211-231. https://doi.org/10.1108/APJML-06-2013-0061

[35] Minavand, H., Minaei, V., Mokhtari, S. E., Izadian, N., & Jamshidian, A. (2013). Project Managers Vs Operations Managers: A comparison based on the style of leadership. *IOSR Journal of Business and Management (IOSR-JBM)*, *12*(5), 56-61. https://doi.org/10.9790/487X-1255661

[36] Farzad, G., & Roshdieh, N. (2024). The Interplay of Destructive Work Behaviors, Organizational Citizenship Behaviors, and Fiscal Decentralization: Implications for Economic Development in Developing Countries.

International Research Journal of Economics and Management Studies IRJEMS, *3*(8), 1-8. https://doi.org/10.56472/25835238/IRJEMS-V3I8P101

[37] Nazari-Shirkouhi, S., & Keramati, A. (2017). Modeling customer satisfaction with new product design using a flexible fuzzy regression-data envelopment analysis algorithm. *Applied Mathematical Modelling*, 50, 755-771. https://doi.org/10.1016/j.apm.2017.01.020

[38] Nazari-Shirkouhi, S., Keramati, A., & Rezaie, K. (2013). Improvement of customers' satisfaction with new product design using an adaptive neuro-fuzzy inference systems approach. *Neural Computing and Applications, 23*, 333-343. https://doi.org/10.1007/s00521-013-1431-x

[39] Nazari-Shirkouhi, S., Keramati, A., & Rezaie, K. (2015). Investigating the effects of customer relationship management and supplier relationship management on new product development. *Tehnički vjesnik, 22*(1), 191-200. https://doi.org/10.17559/TV-20140623130536

[40] Talebzadeh, H., Fattahiamin, A., Talebzadeh, M., Sanaei, F., Moghaddam, P. K., & Espahbod, S. (2024). Optimizing Supply Chains: A Grey-DEMATEL Approach to Implementing LARG Framework. *Tehnički glasnik, 19*(3), 382-389. https://doi.org/10.31803/tg-20240302201341

[41] Gholami, M. H., Asli, M. N., Nazari-Shirkouhi, S., & Noruzy, A. (2013). Investigating the influence of knowledge management practices on organizational performance: an empirical study. *Acta Polytechnica Hungarica, 10*(2), 205-216. https://doi.org/10.12700/APH.10.02.2013.2.14

[42] Ng, B. Y., & Rahim, M. (2005). A socio-behavioral study of home computer users' intention to practice security. *PACIS 2005 Proceedings*, 20. https://aisel.aisnet.org/pacis2005/20

[43] Haeussinger, F., & Kranz, J. (2013). Information security awareness: its antecedents and mediating effects on security compliant behavior. *ICIS 2013 Proceedings, 9*. https://aisel.aisnet.org/icis2013/proceedings/SecurityOfIS/9

[44] Espahbod, S., Tashakkori, A., Mohsenibeigzadeh, M., Zarei, M., Arani, G. G., Dzikuć, M., & Dzikuć, M. (2024). Blockchain-Driven Supply Chain Analytics and Sustainable Performance: Analysis Using PLS-SEM and ANFIS. *Sustainability, 16*(15), 6469. https://doi.org/10.3390/su16156469

[45] Chin, W. W. (1988). Issues and opinion on structural equation modelling. *MIS Quarterly, 22*(1), vii–xvi.

[46] Anbari, M., Talebzadeh, H., Talebzadeh, M., Fattahiamin, A., Haghighatjoo, M., & Jafari, A. M. (2024). Understanding the drivers of adoption for Blockchain-enabled intelligent transportation systems. *Tehnički glasnik, 18*(4), 598-608. https://doi.org/10.31803/tg-20240411223559

[47] Azimi Asmaroud, S. (2022). Preservice Elementary Teachers' Categorical Reasoning and Knowledge Transfer on Definition Tasks with Two Dimensional Figures. *Theses and Dissertations*. 1588. https://ir.library.illinoisstate.edu/etd/1588

[48] Latifi, K., Ebrahimi, A., Ranjbaran, M., Mirzaei, A., & Fakhri, Z. (2023). Efficient customer relationship management systems for online retailing: The investigation of the influential factors. *Journal of Management & Organization, 29*(4), 763-798. https://doi.org/10.1017/jmo.2022.65

Authors' contacts:

Somayye Nasiri
Department of Computer Engineering Zanjan Branch,
Islamic Azad University,
Daneshjoo Blvd., Lower Mountain Lands Road, Zanjan, Iran

Shahram Shahabi
Faculty of Management, North Tehran Branch,
Islamic Azad University,
Vafadar Blvd., Shahid Sadoughi St., Hakimiyeh Exit, Shahid Babaee Highway,
Tehran, Iran

Aynaz Shafiesabet
Department of Finance, Bauer Collage of Business,
University of Houston,
4250 Martin Luther King Blvd Ste 334, Houston, TX 77204, USA

Maral Talebbeidokhti
(Corresponding Author)
Department of Cybersecurity and Networks,
Tagliatela College of Engineering, University of New Haven,
300 Boston Post Rd, West Haven, CT 06516, USA
talebbeidokhtimaral@gmail.com

Elmira Alvandi Behineh
Department of Business Administration,
University of the Cumberlands,
6191 College Station Dr, Williamsburg, KY 40769, USA

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240704100407

Original scientific paper
Received: 2024-07-04, Accepted: 2025-04-29

Accelerating Game Level Design with Machine Learning: A Unity Module for Procedural Kitchen Generation

Petar Pejić*, Jelena Pejić

Abstract: The video game industry faces significant challenges in content creation, with AAA games requiring extensive time and resources. This research addresses these challenges through the development of a Unity module for automatic 3D kitchen model generation using a Machine Learning-based Procedural Kitchen Generation (PKG) model. The module significantly reduces the time needed for game-level design, achieving designs over five times faster than traditional methods. A comparative study shows that the module produces results comparable to industry-standard tools in terms of user preference. This Unity module offers promising potential for commercial use, streamlining the design process and enhancing efficiency in game development.

Keywords: 3D Kitchen Models; Game Level Design; Machine Learning; Procedural Content Generation; Unity Module

1 INTRODUCTION

Computer games are increasingly present in our lives. Every day, hundreds of millions of players around the world are entertained by games [1]. It is expected that the total number of active video gamers worldwide reach 3.02 billion by 2029 [2]. Digital video games represent a large industry, generating sales of around $282.3 Billion worldwide in 2024, with annual growth rates of 8.76% between 2024 and 2027, resulting in a projected market volume of $363.2 Billion by 2027 [3].

AAA games typically take 18-36 months to develop and require the work of a few hundred people, including artists, designers, programmers, and audio engineers [4]. The cost of one AAA game development can exceed $500 million, including marketing budget [5]. The cost breakdown of an average AAA game shows that 40% of the budget is spent on marketing and 37% on generating content [6]. Since 2005, one of the significant challenges, besides cost, has been the difficulty in finding qualified individuals, even within the global workforce [7]. As a consequence, content production has grown to the point at which it has become a bottleneck in both game budgets and product time-to-market [8].

Solutions for this problem can be found in Procedural Content Generation [8, 9, 10, 11] which could eliminate the necessity for a human designer or artist to create that content. Humans are costly and slow, and there appears to be an increasing demand for more of them continuously [12]. Procedural Content Generation can be defined as the algorithmic generation of any content, with no, or limited human intervention. However, it is mostly applied to games, such as rules and mechanics [13, 14], levels [9, 15], music [16], vegetation [17], buildings [11, 17], terrains [19], story [20], and puzzles [21], for instance.

Level generation is one of the oldest and most popular challenges in the PCG domain [22]. Levels are essential content, as every game includes some form of spatial representation or virtual world where the player can engage in various actions [12].

Given their numerous similarities, the procedural generation of game levels can be approached through the perspective of procedural architecture [12]. Much like architecture, game level design must account for two key properties when creating virtual environments: aesthetics and functionality. The aesthetic properties of created virtual environments (geometry, texture, lights, etc.) may have a significant impact on the reality of the game. Besides visual impression, aesthetic also affect the navigation ability, for instance, a series of identical rooms can easily confuse the player. Functional properties are important for the generation of realistic virtual environment. They assume adherence to specific domain knowledge and principles during game level generation: urban design principles in the case of a city, architecture principles in the case of building generation, interior design in a case of furniture layout organization.

For the purpose of this research, we developed a Unity game engine module for automatic linear kitchen 3D model generation. While our earlier work [23] introduced the PKG model and demonstrated its effectiveness, the present study significantly extends that research by embedding the PKG model into the Unity engine, enabling real-time 3D kitchen generation and visualization. To evaluate the Unity module, we conducted a comparative analysis of model creation time between professional game level designers using standard industry tools and our module. Additionally, a perceptual study was performed using a subjective two-alternative forced-choice (2AFC) method, where participants compared kitchens generated by our Unity module to those designed by professionals. The results demonstrate that our system significantly accelerates the 3D modeling process while maintaining a high level of visual and functional quality.

The primary contribution of this research is the development and evaluation of a Unity-integrated module for automatic linear kitchen 3D model generation based on the PKG model. In contrast to the previous study that focused only on layout prediction, our new module offers full automation of 3D model generation inside an industry-standard game engine. Furthermore, the proposed solution is more than five times faster than traditional industry workflows and allows direct real-time integration into applications such as games, VR, and interior visualization tools.

2 RELATED WORKS

Regarding Procedural Content Generation, there exists an overwhelming amount of research papers [1, 10]. Most of them are applied to games' rules and mechanics [14]. Some of them address procedural level generation [22]. One group of research are addresses more specific area of game level design such as generation of buildings [11]. Another group of researchers addresses indoor space generation [24, 25].

There are several studies addressing even more specific problem inside the area of procedural interior generation, which is automatic layout generation. One group of authors [25, 26], uses rule based parametric principles for automatic layout creation. Another group of researchers [24, 23] is oriented toward ML-based algorithms for automatic layout generation.

Only a small number of studies [25, 27] address automatic kitchen layout generation, providing solutions with certain limitations. Several researchers developed a simple system for parametric 3D modeling specifically tailored for I-shaped kitchens [28]. This system was based on an idealistic interior space, disregarding the position of installations such as water pipes and canalization. Another study addresses the same problem by utilizing rule-based algorithms for automatic kitchen layout design, considering the placement of installations [25]. Due to the complexity and limitations of the approach used, the final system achieves only 69.2% accuracy in layout generation. Recent research presents an innovative Procedural Kitchen Generation (PKG) model for linear kitchen layout design, which utilizes data from existing layouts with the help of an advanced machine learning algorithm [23]. The developed PKG model, achieving an accuracy of 77.4%, surpasses the performance of the rule-based system. Consequently, we selected this model for generating layouts in our Unity module.

3 MATERIAL AND METHODS

3.1 PKG Model

The PKG model [23] serves as the core logic for our Unity module, utilizing a pipeline of machine learning classifiers for automatic linear kitchen layout generation. The model's input features include:

- Kitchen total length (mandatory),
- Ventilation position (optional),
- Sink drain position (optional),
- Window position (optional),
- Kitchen entry door position (optional), and
- Exit door position (optional).

The PKG model focuses on spatial and functional characteristics of kitchen elements, such as dimensions, placement constraints, and functional groupings. Visual properties like texture and color were intentionally excluded from the model's input features. This is because the primary objective of the PKG model is to generate structurally valid and functionally realistic layouts. Texture and color, while important for the player's visual experience, are handled separately during the rendering phase in Unity, where materials and appearances are applied to the 3D models. Therefore, the visual realism is achieved in the final visualization, not during the layout generation process.

The model outputs the types and positions of the three major elements:

- Washing (one sink),
- Cooking (one stove), and
- Storage (one refrigerator).

3.2 Unity Module

Game-level designers indicated that, in their practice, all input features except for kitchen length are redundant. The primary concern is the player's visual experience, which should convey the impression of a realistic interior [22]. For that reason, the developed Unity module uses the PKG model with only one input feature – kitchen total length. Using Unity game engine version 2019.3.4f1, a module for the automatic generation and presentation of 3D kitchen models was developed. The module pipeline is presented in Fig. 1.

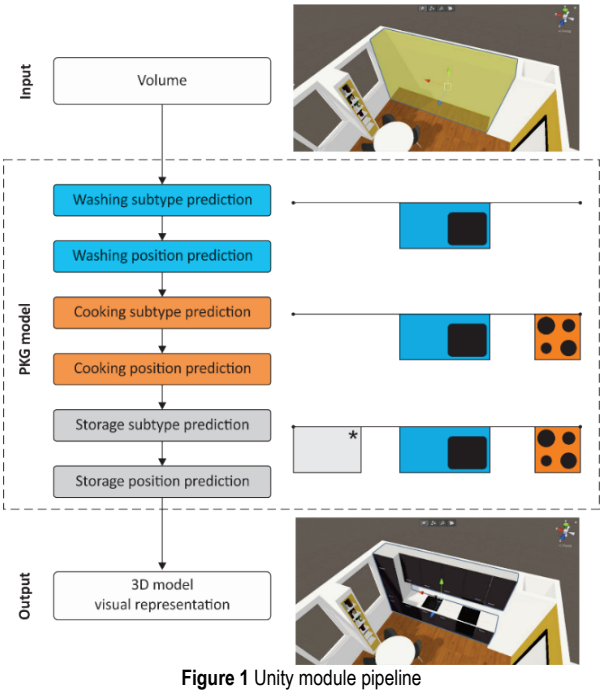


Figure 1 Unity module pipeline

The created module requires manual input of parameters inside the Unity game scene. Marking of kitchen space is done by positioning, rotating, and scaling the cube (Fig. 2a). Users can choose a kitchen design style and can see a preview of the final kitchen 3D model (Fig. 2b).

3D models of kitchen elements for the five most popular kitchen styles [30] were created using Autodesk Maya 2019 and utilized for spatial kitchen representation. These models encompass various subtypes and their respective lengths, including washing, cooking, storage, and additional working surfaces. A total of 50 kitchen element 3D models were produced, with 10 models per style. Models were exported in *.FBX file format for future use and converted inside Unity to final 3D prefabs. Their position and used subtype are determined using a service for communication with the PKG model.

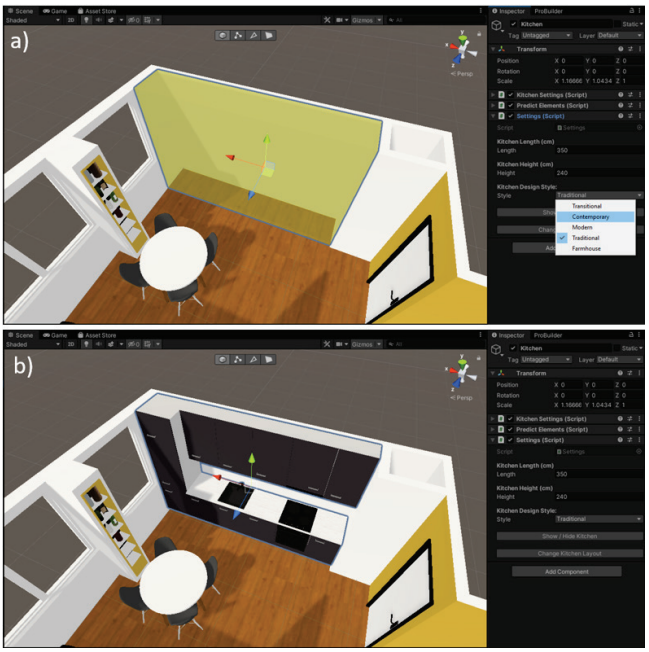


Figure 2 Unity module: a) marking of kitchen space and b) final 3D kitchen model

4 EFFICIENCY STUDY

We conducted an observational efficiency study of our Unity module by recording the time needed for professional game-level designers to create a kitchen 3D model, as well as collecting their subjective evaluations. The primary aim of this study was to explore practical usability and time savings when using our Unity module based on the PKG ML model compared to traditional pipelines.

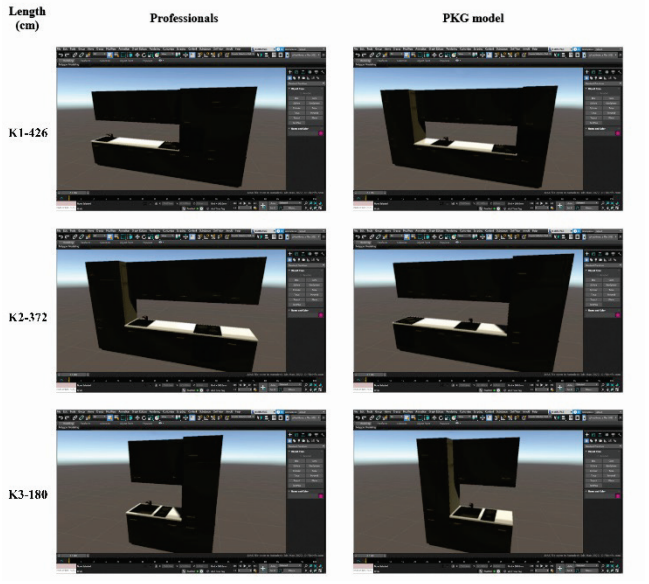


Figure 3 Final 3D models

Eight professional game-level designers participated, including six males and two females, aged 26 to 42, with an average of 11.87 years of experience in the game industry. Although the sample size is relatively small, it was composed of highly experienced professionals to ensure the credibility of the findings within the target application domain. Each participant created three different linear kitchen 3D models (lengths: K1 – 426 cm, K2 – 372 cm, and K3 – 180 cm) first using their preferred pipeline from everyday practice, and then using our Unity module. As input, they received kitchen lengths and kitchen elements in *.FBX format and as Unity prefabs, and they were instructed to present the final designs inside the Unity game engine.

Our study was performed on the same PC (Intel I7/16GB/ssd250+1T/2060gtx) equipped with all necessary software requested by game-level designers. First each of them was asked to create 3D models of three different lengths of linear kitchens using the preferred pipeline with the final result presented inside Unity. After that, they received basic instructions for Unity module use and repeated the process for the same kitchen lengths. Time was measured all 6 times and final 3D models are recorded (Fig. 3). In addition, we asked the game-level designers to give their subjective opinion regarding our Unity module and its usability in their professional work.

5 PERCEPTUAL STUDY

The perceptual experiment was conducted solely using a subjective, two-alternative forced-choice preference approach, where participants were instructed to select the image from each pair that they believed depicted a more plausible arrangement of objects in the room. Each participant completed 27 comparison tasks. 24 kitchen designs created in the "Efficiency study" (Fig. 4a), and 3 "vigilance tests" (Fig. 4b) with comparison featuring obviously wrong answers. We collected responses for each comparison from 32 users who participated in our perceptual study, discarding any responses from individuals who did not achieve 100% accuracy on the vigilance test. The test group consisted of 22 males and 10 females, aged 16 to 44.

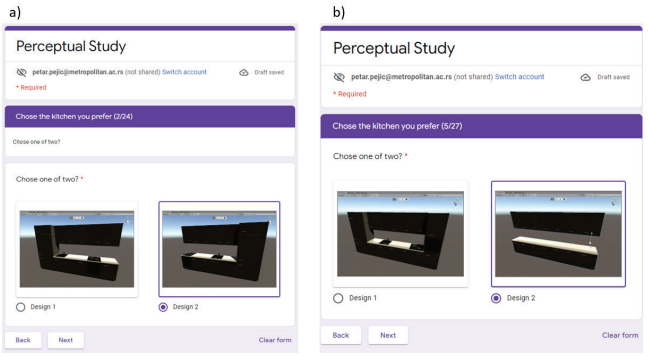


Figure 4 Perceptual study: a) efficiency study and b) vigilance tests

We aimed to determine if there was a significant difference in user preference between kitchen 3D models generated by our Unity module based on ML algorithms and those created using other game industry-standard tools and pipelines. If we found no statistically significant differences between the two methods or if our results were preferred by the users, our method could be considered successful. To validate this, we conducted an experiment using a subjective, two-alternative, forced-choice preference approach, similar to methods used in reference papers [30, 31, 32].

We conducted our user study through an online questionnaire using Google Forms. Each question presented two kitchen designs side by side for the user to compare: one created with our Unity module and the other with game industry-standard tools and pipelines. Each kitchen design was represented by a first-person perspective. Users were required to choose which of the two interior designs they would prefer to see in a game for each question. The questionnaire comprised 27 questions, with both the order of the questions and the placement (left or right) of the two compared interior designs being randomized.

6 RESULTS AND DISCUSSION

The efficiency study measured the time needed for completion of the task for each of the 8 experts (Tab. 1). Depending on their experience and preferences they chose to create a 3D kitchen model using the different pipelines. Two of them used kitchen elements directly inside Unity to create a kitchen 3D model. The rest of them used 3D modeling software to create the kitchen and then import the final 3D model into Unity. Results show significant time saving for each kitchen type using our Unity module based on the PKG ML model. In total, our module is 5.63 times faster than traditional tools and pipelines.

Table 1 Time needed for task completion (mm:ss)

		Traditional pipeline			Our Unity module		
No	Pipeline	K1	K2	K3	K1	K2	K3
1	Unity	01:32	01:52	01:59	00:24	00:22	00:22
2	Unity	01:52	02:10	02:25	00:20	00:19	00:22
3	Maya + Unity	01:52	02:10	02:25	00:28	00:26	00:28
4	3D Max + Unity	02:52	03:10	03:25	00:45	00:48	00:38
5	Maya + Unity	02:22	03:08	03:22	00:18	00:19	00:17
6	Maya + Unity	01:52	02:10	02:25	00:21	00:25	00:25
7	3D Max + Unity	02:52	03:15	03:25	00:38	00:36	00:37
8	Maya + Unity	02:02	02:53	03:21	00:20	00:20	00:19
Average		**02:09**	**02:36**	**02:51**	**00:27**	**00:27**	**00:26**
Total average			**02:32**			**00:27**	

In addition, we asked the game-level designers to give their subjective opinion regarding our Unity module and its usability in their professional work. They reported that our Unity module can be very useful for their work in the case of levels with a lot of interiors. Experts stressed that an important feature for the module to be applicable in a commercial project is a bigger kitchen elements database and the possibility for manual addition of custom-made kitchen elements. A minority of experts expressed interest in Maya and 3D Max plugins for automatic kitchen generation based on the PKG model. While some of them notice that the development of such plugins can be very helpful not only for video game developers but also for interior designers.

To evaluate the perceptual study, we employed a Chi-square nonparametric analysis, in order to identify any statistical significance across our three conditions (three different kitchens). A one-dimensional Chi-square analysis was performed separately for each condition. Each condition included 24 preference answers, which were compared against an expected frequency of 12 results. We computed the Chi-square values and tested them for significance. The results of this analysis are presented in Tab. 2.

Table 2 Perceptual study results

Kitchen	X^2-value	p-value
K1	0.302	0.583
K2	3.191	0.074
K3	0.019	0.891

Furthermore, the recorded frequencies of user preferences are presented in Fig. 5. The Chi-square analysis results indicate that there is no winner, amongst our Unity module and traditional pipelines, which would significantly outperform the other. In summary, our module achieved a higher frequency of user preference than traditional pipelines in all three conditions, but no one of them was statistically significant.

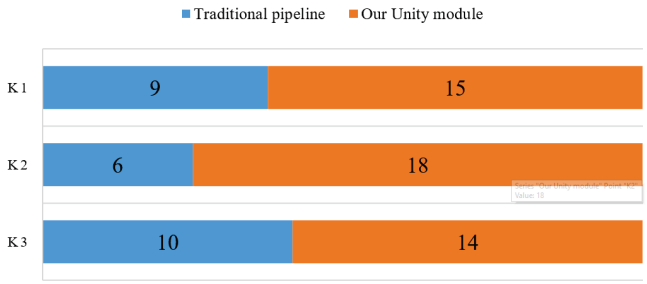


Figure 5 Frequencies of user preferences

7 CONCLUSION

The results of the efficiency and perceptual study demonstrate that our method can produce practical and livable kitchen designs that are comparable to the game industry-standard tools and pipelines. Moreover, the performance measurements show that the use of our Unity module, on average can speed up the process of 3D kitchen model creation by more than 5 times.

Using our Unity module experts need on average 27 seconds to generate a kitchen 3D model, where most of the time is spent on kitchen volumetric input and negligible time on kitchen generation using the PKG model. The fast speed of the PKG model can enable the interactive generation of VR environments during runtime, which have the potential to significantly lower the size of game levels without losing the quality of 3D models.

We find especially valuable experts' subjective opinions regarding our Unity module and its usability in their professional work. All of them found our Unity module based on the PKG model useful with potential for commercial use. The missing features they pointed out, necessary for commercial use are at the same time the tasks for future work and improvement of our Unity module.

Future work on the Unity module should firstly address the creation of a bigger kitchen elements database. Secondly, it should be upgraded to support the user-friendly addition of custom kitchen 3D elements to the database. Additionally, expanding the efficiency study with a larger and more diverse group of participants could further validate and generalize the observed benefits of the Unity module. An alternative direction of future work can be the development of a new plugin based on the PKG model for 3D modeling software such as 3D Max, Maya, and Cinema 4D.

8 REFERENCES

[1] Hendrikx, M., Meijer, S., Van der Velden, J., & Losup, A. (2013). Procedural Content Generation for Games: A Survey. *ACM Transcripts on Multimedia Computing, Communications and Applications*, 1-22. https://doi.org/10.1145/2422956.2422957

[2] Clement, J. (2024). *Number of video game users worldwide from 2019 to 2029.* New York: Statista Market Insights.

[3] Statista Inc. (2023). *Video Games: market data & analysis.* New York: Statista Market Insights.

[4] Walton, M. (2020). *Professional Techniques for Video Game Writing.* Boca Raton: CRC Press.

[5] Ivanov, M., Wittenzellner, H., & Wardaszko, M. (2019). Video Game Monetization Mechanisms in Triple A (AAA) Video Games. *Simulation Gaming Through Times and Disciplines. ISAGA 2019*, 389-404. https://doi.org/10.1007/978-3-030-72132-9_33

[6] Senchenko, A., Patterson, N., Samuel, H., & Ispir, D. (2022). SUPERNOVA: Automating Test Selection and Defect Prevention in AAA Video Games Using Risk Based Testing and Machine Learning. *2022 IEEE Conference on Software Testing, Verification and Validation (ICST).* Valencia. https://doi.org/10.1109/ICST53961.2022.00043

[7] Krueger, D., Brand, O., & Burton, D. (2005). Reinventing your company without reinventing the wheel. *Proceedings of the Game Developers Conference.*

[8] Joen, H., et al. (2023). RaidEnv: Exploring New Challenges in Automated Content Balancing for Boss Raid Games. *IEEE Transactions on Games*, 1-14.

[9] Risi, S., & Togelius, J. (2020). Increasing generality in machine learning through procedural content generation. *Nature Machine Intelligence*, 428-436. https://doi.org/10.1038/s42256-020-0208-z

[10] Amato, A. (2017). Procedural Content Generation in the Game Industry. In: Korn, O., & Lee, N. (eds). In *Game Dynamics* (pp. 15-25). Springer. https://doi.org/10.1007/978-3-319-53088-8_2

[11] Župan, R., Župan, K., Frangeš, S., & Vinković, A. (2018). Automatic procedural 3D modelling of buildings. *Tehnički glasnik, 12*(3), 166-173. https://doi.org/10.31803/tg-20180713143638

[12] Yannakakis, G., & Togelius, J. (2018). *Artificial Intelligence and Games.* Springer. https://doi.org/10.1007/978-3-319-63519-4

[13] Hom, V., & Marks, J. (2007). Automatic Design of Balanced Board Games. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 3*(1), 25-30. https://doi.org/10.1609/aiide.v3i1.18777

[14] Ruiz-Quiñones, J., & Fernández-Lei, A. (2023). Automated Videogame Mechanics Generation with XVGDL. *ICGA Journal*, 124-152. https://doi.org/10.3233/ICG-230227

[15] Morris, J. (2023). A Procedural Generation Platform to Create Randomized Gaming Maps using 2D Model and Machine Learning. *11th International Conference on Signal Image Processing and Multimedia (SIPM 2023).* Vancouver.

[16] Scirea, M., Eklund, P., Togelius, J., & Risi, S. (2018). Evolvingin-game mood - expressive music with Meta compose. *Proceedings of Audio Mostly (AM2018).* https://doi.org/10.1145/3243274.3243292

[17] Garifullin, A., Shcherbakov, A., & Frolov, V. (2022). Fitting Parameters for Procedural Plant Generation. *Computer Science Research Notes*, 282-288. https://doi.org/10.24132/CSRN.3201.35

[18] Rodrigues, F., Neto, J., & Vidal, C. (2015). Split grammar evolution for procedural modeling of virtual buildings. *XVII Symposium on Virtual and Augmented Reality.* Sao Paulo, Brazil. https://doi.org/10.1109/SVR.2015.18

[19] Ahmed, S., & Pandey, B. (2023). Procedural Terrain Generation by Sampling a 2D Monochrom Perlin Noise Map in Unity. *Asian Journal of Research in Computer Science*, 37-42. https://doi.org/10.9734/ajrcos/2023/v16i1333

[20] Cheong, Y., & Young, M. (2015). Suspenser: A Story Generation System for Suspense. *IEEE Trans. Comput. Intell. AI Games*, 39-52. https://doi.org/10.1109/TCIAIG.2014.2323894

[21] Trenquelleon, G., Choukarah, A., Roucairol , M., Addoum, M., & Cazenave, T. (2023). Procedural Generation of Rush Hour Levels. *Computers and Games. CG 2022. Lecture Notes in Computer Science vol 1386.* https://doi.org/10.1007/978-3-031-34017-8_15

[22] Khalifa, A., Perez-Liebana, D., Lucas, S., & Togelius, J. (2016). General video game level generation. *Genetic and Evolutionary Computation Conference.* https://doi.org/10.1145/2908812.2908920

[23] Pejic, J., & Pejic, P. (2022). Linear kitchen layout design via machine learning. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 1-12. https://doi.org/10.1017/S089006042100038X

[24] Wang, K., Savva, M., Chang, A., & Ritchie, D. (2018). Deep convolutional priors for indoor scene synthesis. *ACM Transactions on Graphics*. https://doi.org/10.1145/3197517.3201362

[25] Pejic, P., Mikic, M., & Milovanovic, J. (2019). Automatic Rule-based Kitchen Layout Design. *PaKSom 2019.*

[26] Stojakovic, V., Budak, I., Obradovic, R., & Koralja, D. (2017). Parametric modeling applied to the virtual reconstruction of the damaged sculpture of St John Nepomuk in Petrovaradin. *eCAADe.* Rome. https://doi.org/10.52842/conf.ecaade.2017.2.389

[27] Henderson, P., Subr, K. & Ferrari, V. (2019). Automatic Generation of Constrained Furniture Layouts. *Computer Vision and Pattern Recognition*.

[28] Pejic, P., Jovanovic, D., Marinkovic, J., & Krasic, S. (2019). Parametric 3D modeling of I-shape kitchen. *JIDEG*, 155-158.

[29] Billock, J. (2024. 5. 29.). *The 16 Most On-Trend Kitchen Styles Interior Designers Can't Get Enough Of.* (The kitchen) Retrieved 6. 12. 2024., from https://www.thekitchn.com/kitchen-styles-23661388

[30] Kan, P., & Kaufmann, H. (2018). Automatic Furniture Arrangement Using Greedy Cost Minimization. *Conference on Virtual Reality and 3D User Interfaces (VR).* https://doi.org/10.1109/VR.2018.8448291

[31] Jimenez, J., Sundstedt, V., & Gutierrez, D. (2009). Screen-space perceptual rendering of human skin. *ACM Trans. Appl. Percept*, 1-15. https://doi.org/10.1145/1609967.1609970

Authors' contacts:

Petar Pejic, Associate Professor
(Corresponding author)
Faculty of Information Technology, Belgrade Metropolitan University,
Tadeuša Košćuška 63, 11158 Belgrade, Serbia
+381 11 203 0885, petar.pejic@metropolitan.ac.rs

Jelena Pejic, Teaching Assistant
Faculty of Sciences and Mathematics, University of Niš,
Višegradska 33, 18106 Niš, Serbia
+381 18 533015, jelena.pejic@pmf.edu.rs

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240806064110

Original scientific paper
Received: 2024-08-06, Accepted: 2025-01-20

Customer Relationship Management Processes and Practices in a Higher Education Institution: A Case Study

Samuel Lakkala*, Joonas Suutari, Pekka Tervonen, Harri Haapasalo

Abstract: This study examines the implementation and advancement of customer relationship management (CRM) systems in higher education institutions (HEIs) through a detailed case study. It aims to (1) understand CRM processes from an HEI perspective and (2) develop CRM processes and practices tailored to HEI needs. The results highlight the critical role of CRM in managing diverse stakeholder relationships amidst growing competition for students, partners and funding. Essential processes, such as partnership creation, university-industry collaboration and systematic activity logging, are emphasised for strategic CRM application. Key CRM practices identified include process mapping, standardised procedures and clear user roles. The study provides a framework for refining CRM strategies in HEIs, enhancing stakeholder engagement and operational efficiency. These findings offer HEIs valuable insights for optimising CRM systems and fostering improved CRM in a competitive educational landscape.

Keywords: CRM implementation; CRM processes; CRM systems; customer relationship management; higher education

1 INTRODUCTION

Customer relationship management (CRM) and CRM systems have traditionally been part of the toolbox of business life, but nowadays educational and research organisations are also and increasingly adopting CRM systems and customer-oriented methods of operation. The competition between universities for students, partners and funding has also boosted the need to develop and systematise CRM in the university world [1].

CRM is usually strongly associated with information systems as information systems help in collecting, storing and analysing customer data [2]. However, Payne and Frow [3] emphasise that CRM should always be seen broadly rather than just as a technology solution if a CRM system project is to be successful. In connection with the system's implementation, the necessary processes that fit the organisation's CRM goals must be defined [4].

CRM systems are extensively utilised in the business sector, supported by a substantial body of research focused on their application within this context. In contrast, research on CRM in HEIs is still in its nascent stages [5]. The majority of existing studies have viewed students as the primary customers for HEIs, as highlighted by [6-9]. Meanwhile, research addressing other significant stakeholder groups, such as alumni and the broader business and industrial community, remains relatively underdeveloped [10, 11]. This gap underscores the need for further studies on the diverse applications and effects of CRM systems in HEIs that go beyond the predominant student-centric perspective.

To address this research gap, the present work presents a case study focusing on the implementation of a CRM system at a Finnish public university. The goal of this investigation was to gain new insights into the role and application of CRM systems and their processes within an educational setting, moving away from the traditional focus on students. The study was guided by the following research questions:

RQ1: What is CRM and its processes from the perspective of an HEI?

RQ2: How are the CRM processes and practices of a HEI created in connection with the implementation of a CRM system?

To address these research questions, a qualitative research approach was employed, encompassing a literature review and a case study. The literature review that answers to the first research question aimed to establish a theoretical foundation and provide insights into CRM systems within HEIs and related processes. To find an answer to the second research question, a case study was conducted to investigate the implementation of a CRM system and development of processes within the case organisation.

2 LITERATURE REVIEW

Buttle and Malkian [2] define CRM as a core business strategy that integrates internal processes and functions with external networks to create and deliver value to targeted customers profitably. This strategy relies on high-quality customer data and is facilitated by information technology. The effectiveness of a CRM strategy depends on balancing three key resources: people, technology, and processes [3]. Efficient resource utilisation in CRM leads to high customer satisfaction, retention, and loyalty [12]. While CRM systems are widely used in the business sector and extensively studied, there is limited research on their application within HEIs [5].

2.1 Customer Relationship Management in Higher Education Institutions

The implementation of CRM systems is becoming strategically necessary for HEIs [5]. HEIs themselves are no different from other organisations insofar as they are for-profit structures that produce education and research for money [2]. Large HEIs, such as universities and universities of applied sciences, have numerous stakeholders, including students, academic and non-academic personnel, collaborating companies and researchers, various authorities

and other public administration organisations and universities [5, 13, 14].

CRM is typically closely linked with information systems, which aid in the collection, storage and analysis of customer data [2]. These systems gather information about both internal and external customers and consolidate it into a single database for processing. Initially developed for business use, CRM systems have evolved to manage the complex relationships within HEIs. By integrating customer information from various sources, they have become essential for decision-making and managing diverse stakeholder relationships [2, 15].

Rirchards and Jones [16] outline several core benefits of CRM systems for organisations: identifying profitable customers, integrating offerings across channels, improving sales efficiency and effectiveness, enabling personalised marketing, customising products and services, enhancing the efficiency and effectiveness of customer service, and optimising pricing strategies. In the context of higher education, CRM systems not only support the monitoring of students' academic progress and the recruitment of new students but also encourage the development of donor, industry and alumni relations. The systematic management of these relationships, facilitated by CRM's data consolidation, communication tools and event management capabilities, empowers HEIs in their daily operations and strategic initiatives, such as funding applications and multi-channel communication [1, 10, 14].

Despite their clear benefits, the implementation of CRM systems in HEIs faces significant challenges. A major one is possible resistance from academic staff toward adopting business-oriented systems and methodologies, which underscores the internal cultural barriers to CRM projects in academic settings [6]. This resistance is heightened by the increasing expectations for academics to participate in external research, contract research, consultancy and executive education [11]. Additional challenges, such as the lack of a unified CRM strategy, inadequate change management and insufficient understanding of existing processes, further complicate CRM system implementation. Adapting commercial CRM systems, originally developed for business environments, to address the specific needs of HEIs also presents significant difficulties [1].

2.2 CRM Processes in HEIs

The foundation of effective CRM implementation lies in well-defined processes and practices. It encompasses the entirety of how customer information is gathered, managed and utilised to foster relationships. The establishment of clear processes ensures consistency and efficiency across the organisation, aligning with the strategic goals of CRM implementation [3, 12, 14, 17].

The CRM implementation model created by [12], shown in Fig. 1, combines the three most central dimensions of CRM (people, processes and technology) into an organisation-wide, customer-oriented, technologically integrated, multidisciplinary entity. Thus, the success of the implementation depends on the creation of a suitable combination of the three dimensions [12].

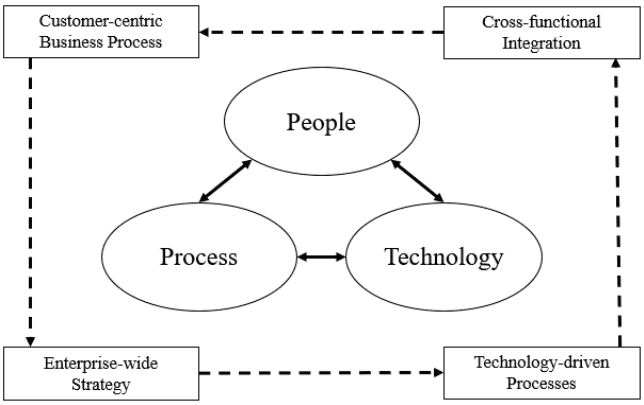


Figure 1 CRM implementation model [12].

Reinartz et al. [18] characterise CRM as fundamentally process-based, highlighting the importance of systematically and proactively managing customer relationships from their initiation to conclusion. These relationships evolve through various stages, each requiring specific management strategies. Effective CRM coordination relies on well-defined processes to meet organisational needs. Developing CRM processes necessitates a customer-oriented approach that begins with identifying the organisation's needs and translating them into CRM requirements. This involves automating decision-making, redesigning processes, setting precise and measurable implementation goals, enhancing stakeholders' technical capabilities and fostering innovation [3, 12, 14, 17].

Organisations must define CRM processes aligned with their CRM goals and allocate sufficient resources to support them. The suitability of these processes is influenced by the organisational structure and culture. It is crucial to ensure that processes function as intended, supported by the necessary technology, which is typically a CRM system designed around these processes [4]. During CRM system implementation, it is both feasible and cost-effective to develop processes and tools simultaneously, ensuring that detailed process definitions serve as support materials, guaranteeing processes are well-defined and supported by the system [19]. Additionally, the role of employees is fundamental to the success of CRM implementation, as they are the end-users of the system. Management must establish clear guidelines for using the CRM system and implementing new processes while ensuring compliance with these guidelines [20].

In the context of CRM in HEIs, university-industry collaboration (UIC) refers to the interaction between universities and industries, aiming to foster the exchange of information and technology. The customer acquisition process in HEIs mirrors the UIC creation process, which initiates collaboration between the university and industry. The complexity and formality of the relationship dictate the number of steps in the UIC creation process [21] (Tab. 1). Key issues in planning customer acquisition include

identifying target customers, approaches to engagement and the offerings provided [2].

Table 1 UIC creation process (modified from [21])

Step	Content
1. Identifying the partnership	- Defining the purpose for partnership - Gathering general knowledge about potential partners - Examining past relationships
2. Contacting	- Contacting the identified potential partners
3. Partner evaluation and selection	- Objectively assessing the strategic interest of potential partners - Analysing the alleged and actual capabilities of potential partners - Defining and organising the mix of suitable partners - Selecting partners.
4. Partnership negotiations	- Determining and agreeing on the goal or mission/vision of the partnership - Defining specific common goals - Defining the organisational structure of the partnership - Defining the management and control of the partnership and the division of responsibilities - Approving the plan - Identifying success metrics - Defining intermediate and final results
5. Signing the contract	- Preparating and signing a cooperation or intellectual property agreement

2.3 Synthesis of the Literature Review

CRM can be characterised a core business strategy integrating internal processes and external networks to deliver value relying on high-quality data and information technology. Its effectiveness hinges on balancing people, technology and processes. Efficient CRM leads to high customer satisfaction, retention and loyalty. Although extensively studied in the business sector, CRM application in HEIs is less explored. HEIs aim to produce education and research profitably, managing numerous stakeholders, including students, staff, companies, and public organisations.

CRM systems collect, store, and analyse customer data, consolidating it for processing. Initially for business use, these systems now manage complex relationships within HEIs, supporting decision-making and stakeholder management.

CRM benefits include identifying profitable customers, integrating offerings, improving sales efficiency, enabling personalised marketing, customising services, enhancing customer service, and optimising pricing. In HEIs, CRM systems help monitor student progress, recruit students, and develop donor and alumni relations, empowering them through data consolidation and improved communication.

Effective CRM implementation relies on well-defined processes ensuring consistency and efficiency aligned with the organisation's strategic goals. The CRM model integrates people, processes and technology, where processes transform resources into effective capabilities. Developing CRM processes requires a customer-oriented approach, automating decisions, redesigning processes, setting measurable goals and enhancing technical capabilities. HEIs must define processes aligned with CRM goals, supported by sufficient resources, considering organisational structure and the academic culture. Detailed process definitions aid implementation, ensuring the processes are well-defined and supported by the system.

With respect to RQ1, in the context of HEIs, CRM systems manage complex relationships with multiple stakeholders, as noted. These systems collect, store and analyse customer data, supporting decision-making and stakeholder management. CRM processes in HEIs involve systematically identifying and engaging stakeholders and facilitating and tracking communications. Additionally, they use data-driven insights to evaluate partnerships and make informed decisions while ensuring continuous engagement with stakeholders.

3 MATERIALS AND METHODS

3.1 Research Process and Methods

This study began with a comprehensive literature review to establish a theoretical foundation and understand the significance of CRM systems in HEIs, as well as the processes of CRM. Building on this foundation, it employed a qualitative approach – specifically, a descriptive case study – to thoroughly investigate the implementation of a CRM system within the faculty of a Finnish multidisciplinary public university. The study originated from an ongoing CRM implementation project within one of the university's faculties. This methodological choice, guided by [22] enabled an in-depth examination of CRM practices within the case organisation, providing the rich and detailed insights characteristic of qualitative research [23]. The research process is shown in Fig. 2.

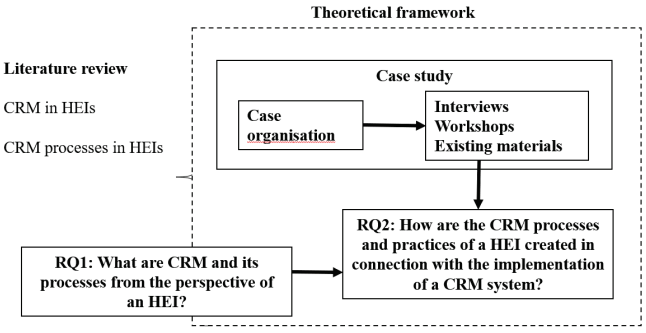


Figure 2 Research process

The research material was collected by interviewing 13 heads of the faculty's research units, observing implementation planning meetings and holding five implementation-related workshops (Tab. 2). The interviews were conducted as part of the project aiming to assess the current state of CRM within the faculty, identify the requirements for the CRM system and determine the critical success factors for its implementation.

Based on the interviews, it was concluded that the initial steps of the project should focus on establishing the primary CRM processes, along with the practices and roles necessary for their execution. Comments and opinions about the current

state of CRM were widely heard from the employees of the pilot project unit during the meetings. The employees were also able to influence the introduction of the system through open events and communication.

Table 2 Sources of research material

Sources of data
Interviews (13), workshops (5), existing materials (university documents and websites) and practical observations (project meetings and trainings)

3.2 Case Organisation

The case organisation was chosen for this research based on several compelling factors. Firstly, the faculty represents a complex entity within a larger university framework, employing 414 staff members, including 33 professors, and serving over 2,300 students across various degree programmes. This environment provides a rich context for examining the intricacies of CRM implementation in higher education.

However, the main reason for selecting this faculty was its active participation in collaborative research projects with a diverse array of external stakeholders, such as companies, civil organisations and governmental bodies. These collaborations, often facilitated by CRM systems, are vital for the faculty's research productivity and its funding, half of which comes from external sources. The faculty's broad network and dependence on external funding highlight the importance of CRM in managing these relationships.

Also, the faculty had recently initiated a CRM implementation project, adding a timely and relevant aspect to the study. Investigating this ongoing project allowed for real-time insights into the challenges and successes of CRM adoption in a higher education setting.

Finally, the involvement of research unit heads in the study was crucial for gaining a comprehensive understanding of CRM's impact at different organisational levels. These leaders had firsthand knowledge of their units' most significant partnerships and projects, providing invaluable insights into the most critical and necessary processes for the organisation.

Fig. 3 illustrates the case organisation's stakeholders and highlights the central role of the CRM system in managing relationships and processes within this ecosystem. At the middle of the organisation are its core stakeholders including students and personnel, who drive the university's daily operations. Surrounding the core, the figure showcases key external stakeholders such as network of industry partners, other research institutions, alumni, donors, funding bodies and event attendees. These groups benefit directly from the CRM system, which facilitates communication and collaboration.

Beyond the immediate stakeholders, the figure expands to include the university's broader external ecosystem, comprising for example, service providers, international partners and governmental bodies. While these stakeholders may not interact directly with the CRM system, they are influenced by the outcomes it enables, such as improved research impact, public outreach and institutional reputation.

The CRM system is portrayed as an overlapping layer within the figure, symbolising its integrative role. It connects internal and external stakeholders, ensuring transparency, consistency, and efficiency in engagement and collaboration.

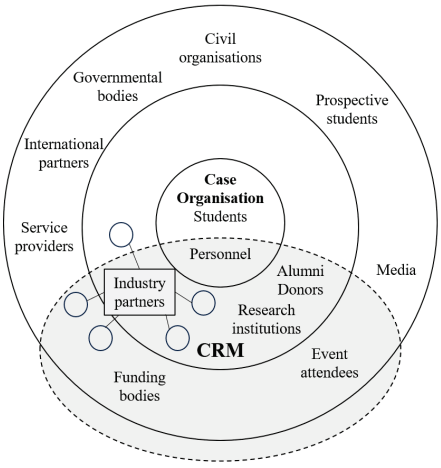


Figure 3 Case organisation's stakeholders

4 RESULTS

Through the interviews, meetings and workshops of the research unit, CRM at the case organisation was defined as 'procedures that develop open, unified and transparent customer and stakeholder activities'. CRM there was based on pre-agreed rules and internal cooperation, and its goal was a positive and professional customer experience in all the university's customer channels. The need for more effective CRM was brought up by emphasising that the CRM system offered a competitive advantage.

It was thought that incorporating the CRM system into the university's existing project management and communication processes would significantly enhance coordination and operational efficiency. Interviews revealed a strong anticipation that this integration would lead to a unified approach in managing projects and increase communication efficacy both internally and with external stakeholders. Such streamlined processes are significantly supported by CRM technologies, such as stakeholder registers and communication tools, optimising interactions and the management of stakeholder information [1].

Based on the interviews and literature, the key CRM processes identified as essential for the initial development of the system were the creation of a partnership, a cooperation process and the logging of activities. The processes described in this section were developed using structured process mapping techniques which enabled the detailed visualisation of workflows, making it easier to identify critical steps, define responsibilities, and ensure alignment with the university's CRM objectives. Process mapping helped visualising the identified processes into manageable components, facilitating clarity and consistency. CRM practices were created to support the processes, which include rules, instructions for operation and user roles including job descriptions.

In the process description terminology, 'CRM user' refers to a person in the research unit who uses the CRM system. A CRM user can be, for example, a professor, researcher, teacher or research assistant, as long as they have a user license. 'CRM system' refers to the CRM system to be implemented. 'Partner' means a cooperating organisation, that is, a company, financier or other organisation, and 'partnership' is a cooperative relationship. 'Company data' includes all information about the company or organisation in the CRM system, including historical data from previous cooperation and company contacts. 'Contacts' are the company's personnel with whom the contact takes place.

4.1 Partnership Creation Process

The partnership creation process is pivotal for establishing formal collaborative agreements between the case organisation and new partners. Utilising a CRM system enhances the efficiency and effectiveness of this process. This section details each stage of the partnership creation process, highlighting the significance of each step and the role of the CRM system in facilitating these interactions. The partnership creation process followed the UIC process as introduced by [21]. Fig. 4 illustrates the process flow.

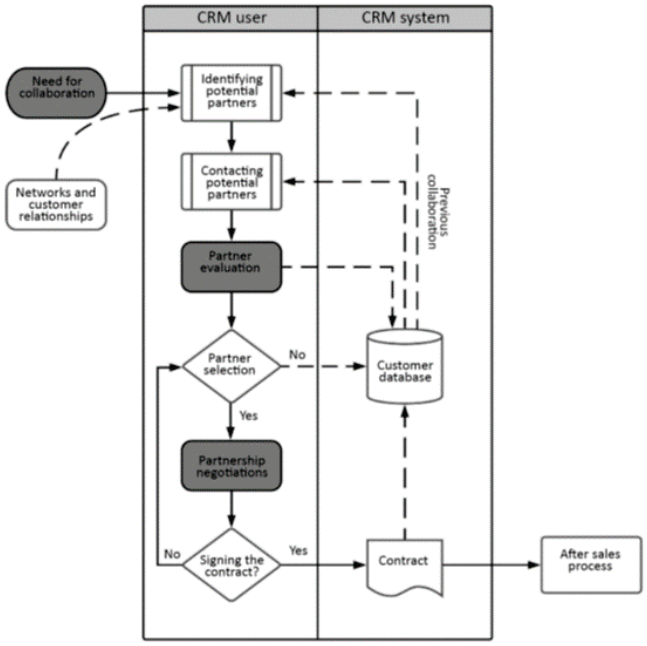


Figure 4 Partnership creation process

Identifying the Need for Collaboration. The initiation of a partnership begins with recognising the need for collaboration. This need can arise from various factors, such as upcoming projects, funding requirements, academic course needs or potential industry collaborations.

Identifying Potential Partners. Once the need for collaboration is established, the next step involves identifying potential partners. This stage includes researching and compiling a list of organisations that align with the goals and requirements of the research unit. The CRM system plays a significant role here by providing a database of existing networks and customer relationships, which can be leveraged to identify suitable partners. Utilising historical data stored in the CRM system ensures that potential partners are evaluated based on previous interactions and collaborations [15].

Contacting Potential Partners. After identifying potential partners, the user initiates contact. This involves reaching out to potential partners through various communication channels, such as emails, phone calls or face-to-face meetings. The CRM system aids by automating communication processes, tracking interactions and scheduling follow-ups [1]. Effective communication is vital to convey the partnership's value proposition and gauge the interest of potential partners.

Partner Evaluation. The evaluation stage is critical for assessing the suitability of potential partners. This includes an assessment of their strategic alignment, capabilities and potential for collaboration. The CRM system facilitates this by providing comprehensive data analysis tools to evaluate the strategic interest and capabilities of potential partners. Evaluating partners ensures that the chosen partners can contribute effectively to the collaborative efforts, thereby enhancing the partnership's overall value [2].

Partner Selection. Based on the evaluation, the most suitable partners are selected. This decision-making process involves comparing the strengths and weaknesses of potential partners and determining which partner best aligns with the research unit's goals. The CRM system supports this stage by providing detailed reports and insights from the evaluation phase, aiding in informed decision-making [2, 17].

Partnership Negotiations. Once a partner is selected, the next step involves negotiating the terms of the partnership. This includes discussing and agreeing the goals, responsibilities and expectations of both parties. The CRM system helps manage these negotiations by maintaining records of the communications and agreements, ensuring that all aspects of the partnership are documented and accessible. Effective negotiations are crucial to the establishment of a clear understanding and agreement between the partners, setting the foundation for a successful collaboration [19].

Signing the Contract. The final stage of the partnership creation process comprises the formalisation of the agreement through a contract. The contract outlines the terms and conditions of the partnership, including roles, responsibilities, deliverables and timelines. The CRM system ensures that the contract details are integrated into the customer database, providing a reference for future interactions and ensuring that all contractual obligations are tracked and met [19]. Formalising the partnership through a contract is essential for legal and operational clarity, ensuring that both parties are committed to agreed-upon terms.

Each stage of the partnership creation process is integral to building an effective collaboration. By systematically identifying, evaluating and selecting partners, organisations can ensure that their collaborations are strategically aligned and mutually beneficial. The CRM system enhances this process by providing a structured approach to managing interactions, maintaining comprehensive records and facilitating data-driven decision-making. This structured approach not only improves the efficiency of the partnership

creation process but also increases the likelihood of successful and sustainable collaborations.

4.2 Cooperation Process

The cooperation process with existing partners is integral to sustaining and enhancing long-term relationships; it closely mirrors the partnership creation process but is generally more streamlined. According to the literature, maintaining and enhancing relationships with existing partners is often more cost-effective than acquiring new ones [2, 24]. Fig. 5 illustrates this process.

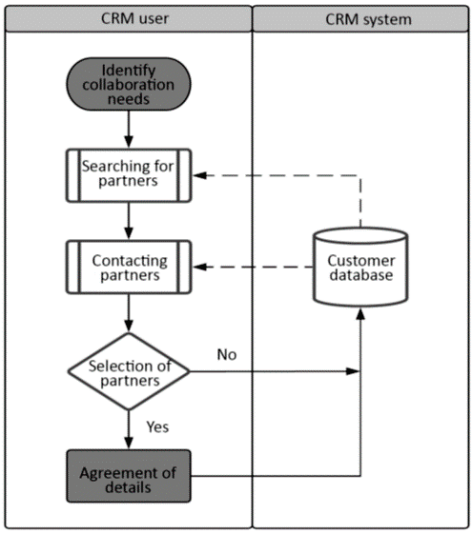


Figure 5 Cooperation process.

By systematically identifying collaboration needs, searching for partners, contacting, evaluating and finalising agreements, organisations can ensure that their cooperative efforts are strategically aligned and mutually beneficial. The CRM system enhances this process by providing a structured approach to managing interactions, maintaining comprehensive records and facilitating data-driven decision-making [4, 18]. Maintaining effective cooperation with existing partners fosters a stable and reliable network of collaborators, contributing to the organisation's strategic goals and enhancing overall performance. This structured approach not only improves the efficiency of the cooperation process but also increases the likelihood of successful and sustainable collaborations.

4.3 Logging of Activities

One of the most critical objectives in implementing a CRM system is the comprehensive recording of all activities with current and potential future partners. Proper documentation of the activities enhances transparency, accountability and the efficient management of relationships. The wide-ranging activities include phone calls, emails, meetings, visits, guest lectures, projects and collaborative work with students [5, 14]. This section elaborates on the activity logging process, its stages and the role of the CRM system in ensuring accurate and effective documentation. Fig. 6 illustrates the activity logging process.

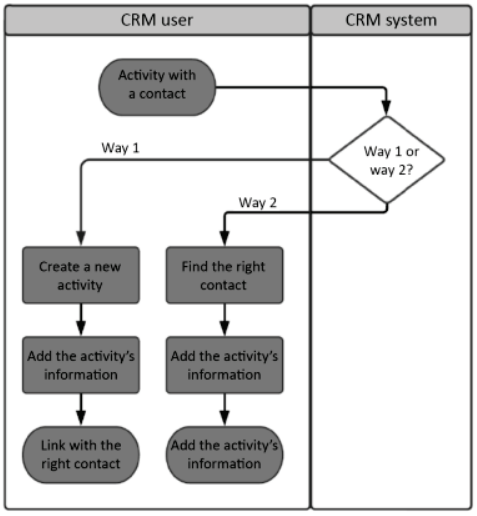


Figure 6 Activity logging process.

The process begins when a CRM user initiates an activity with a contact, such as a scheduled meeting, phone call or project update. The prompt recognition and recording of these activities ensure that all interactions are documented, maintaining a comprehensive history of engagements with partners [15, 19].

At the second stage, the CRM user determines the method of logging the activity, either directly from the start screen or through the specific partner or contact entry. This decision depends on the user's workflow preferences and the nature of the interaction. Both methods provide flexibility and convenience in documenting activities, catering to different user needs and preferences.

When logging directly from the start screen, the user begins by creating a new activity entry in the CRM system, selecting the type of activity and entering preliminary details. They then add specific details, including the date, time, participants and a brief description of the interaction. Detailed documentation at this stage is crucial for future reference and follow-up actions. Finally, the user links the activity to the appropriate contact within the system, ensuring accurate and comprehensive tracking of all interactions [4].

Alternatively, the user may log the activity via the partner or contact entry by first searching for the relevant contact within the CRM system. After navigating to the contact's profile, the user creates a new activity entry linked directly to this contact, ensuring immediate association with the correct partner. As with the direct logging method, the user then enters detailed information about the activity to ensure thorough documentation.

Logging activities accurately and comprehensively is vital for several reasons. First, detailed activity records provide a complete history of all interactions with partners, facilitating better relationship management and informed decision-making [2, 19]. Second, recording activities ensures accountability by documenting who interacted with whom, when and about what, contributing to internal knowledge and evaluation processes. Third, comprehensive activity logs enable CRM users to efficiently follow up on previous interactions, ensuring continuity and coherence in communications. Finally, the aggregated data from logged activities can be analysed to derive insights into engagement

patterns, partner responsiveness and areas for improvement, supporting strategic planning and operational optimisation [2, 15].

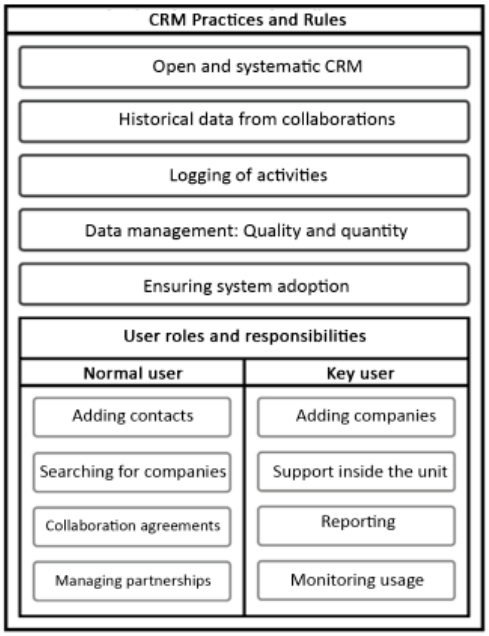


Figure 7 CRM practices and rules in the case organisation.

The practices and rules established in this study aim to ensure the successful implementation of CRM systems in HEIs. By promoting open and systematic CRM usage, these practices make historical information and cooperative activities visible to all personnel within the system. The summary of common practices, rules and user roles in Fig. 7 encapsulates the results and conclusions.

4.4 Practical Implications

The implementation of CRM processes within an HEI is supported by consistent CRM practices and rules. These practices ensure systematic use of the CRM system across the research unit, establishing a framework for effective CRM. This framework includes CRM rules, general instructions for CRM and clear definitions of employee roles and responsibilities.

The CRM practices outlined in this study are crucial for several reasons. Firstly, they help mitigate implementation risks by ensuring data protection and the accuracy of the stakeholder register, even post-implementation. Also, common rules necessitate compliance at the individual level within the research unit, promoting uniformity and accountability. Monitoring user activities, meanwhile – particularly when editing data – ensures adherence to these rules, thereby maintaining data integrity and reliability. Although the practices developed are not exhaustive, they have provided a robust foundation for initial CRM usage in the case organisation.

A key aspect of implementing CRM rules is the systematic recording of all activities in the CRM system. Activities involving external stakeholders should be logged to facilitate tactical and strategic development. Transparency in CRM is a central theme as it enhances cooperation between units, simplifies reporting and promotes multidisciplinary research opportunities. Transparency also aids the formation of larger project groups and collaborations as well as providing staff with a clear understanding of partner interactions.

Performance metrics would further strengthen these practices and processes by enabling a measurable framework for CRM evaluation. During the initial stages of CRM implementation, the focus should be on metrics related to data quality and system usage. For example, organising partners by areas of expertise and maintaining accurate, comprehensive contact information ensures high data quality. Likewise, tracking the number of active users, the availability of support resources, and user feedback offers insights into the adoption of the system and identifies areas requiring improvement. Logging all interactions not only ensures consistency but also provides a unified view of partnerships, avoiding duplications and improving stakeholder experiences. Although financial performance can be more challenging to measure in higher education context, growth in external funding or the establishment of new cooperation agreements can serve as a performance metric.

The ongoing CRM implementation project in the case organisation is planned to be expanded to other faculties in the university. The process descriptions and practices created because of the work outlined here may not be directly suitable for use by other research units, but they might be used as help when describing their own processes and practices. In the CRM implementation performed by other units, the experiences and results obtained from the case organisation of the pilot project will certainly serve at least in a guiding form.

One topic of further research could be the development of CRM in a more strategic direction and the introduction of new functions. The implementation of the pilot project focused on the basic features of the selected CRM system as planned, but the system also offers many other CRM and sales tools. Key account management, project portfolio management, lead generation and sales pipeline and marketing systematisation are examples of possible development directions.

5 CONCLUSION

This article has reported on the implementation of CRM in an HEI and related processes and practices. Bringing a tool from the business world to universities arouses discussion, and its introduction requires effective change management to establish its use. The strategic importance of CRM is emphasised as competition between universities for students, researchers and funding increases.

The first goal of the study was to define CRM and its processes from the perspective of an HEI. CRM means more than just an information system, and the implementation of a CRM system must integrate the aspects of process, technology, people and culture. The special features of an HEI's CRM are strong autonomy, numerous external stakeholders and the coordination of the universities' several tasks prescribed by law. Here, following the initiation of cooperation between the university and industry, the

customer acquisition process of an HEI has been defined from the CRM processes.

The second goal was to create CRM processes and practices related to the implementation of a CRM system suitable for the research unit's operations. The processes focused on creating a new customer relationship and business cooperation, and practices for implementing CRM were created to support the processes. The most important result is an open, up-to-date and systematic recording of cooperation using the system.

6 REFERENCES

[1] Badawood, A. (2016). Critical success factors for implementing customer relation management (CRM) system within university context concepts and literature review the Gulf Region Perspective, *International Journal of Computer Science and Technology*, *7*(1), 125-130.

[2] Buttle, F., & Malkian. S. (2019). *Customer Relationship Management: Concepts and Technologies* (4th ed.). Elsevier. https://doi.org/10.4324/9781351016551

[3] Payne, A., & Frow, P. (2005). A strategic framework for customer relationship management. *Journal of Marketing*, *69*, 167-176. https://doi.org/10.1509/jmkg.2005.69.4.167

[4] Keramati, A., Mehrabi, H., & Mojir, N. (2010). A process-oriented perspective on customer relationship management and organizational performance: An empirical investigation, *Industrial Marketing Management*, *39*(7), 1170-1185. https://doi.org/10.1016/j.indmarman.2010.02.001

[5] Khashab, B., Gulliver, S., & Ayoubi, R. (2020). A framework for customer relationship management strategy orientation support in higher education institutions, *Journal of Strategic Marketing*, *28*(3), 246-265. https://doi.org/10.1080/0965254X.2018.1522363

[6] Daradoumis, T., Rodríguez-Ardura, I., Faulin, J., Juan, A., Xhafa, F., & Martínez-López, F. (2010). CRM applied to higher education: developing an e-monitoring system to improve relationships in e-learning environments, *International Journal of Services Technology and Management*, *14*(1), 103-125. https://doi.org/10.1504/IJSTM.2010.032887

[7] Guilbault, M. (2016). Students as customers in higher education: reframing the debate. *Journal of Marketing for Higher Education, 26*(2), 132-142. https://doi.org/10.1080/08841241.2016.1245234

[8] Wali, A., & Wright, L. (2016). Customer relationship management and service quality: Influences in higher education. *Journal of Customer Behaviour, 15*(1). 67-79. https://doi.org/10.1362/147539216X14594362873532

[9] Octavia, A., Jovanka, D. R., Alqahtani, T. M., Wijaya T. T., & Habibi. A. (2023). Key factors of educational CRM success and institution performance: A SEM analysis, *Cogent Business & Management*, *10*(1). https://doi.org/10.1080/23311975.2023.2196786

[10] Suomi, K., Kuoppakangas, P., & Stenvall, J. (2019). Revisiting "the shotgun wedding of industry and academia"—empirical evidence from Finland. *International Review on Public and Nonprofit Marketing*, *16*, 81-102. https://doi.org/10.1007/s12208-019-00220-8

[11] Evans, N., & A. Miklosik. A. (2023). Driving digital transformation: Addressing the barriers to engagement in university-industry collaboration. *IEEE Access*, *11*, 60142-60152. https://doi.org/10.1109/ACCESS.2023.3281791

[12] Chen, I. J., & Popovich, K. (2003). Understanding customer relationship management (CRM). *Business Process Management Journal*, *9*(5), 672-688. https://doi.org/10.1108/14637150310496758

[13] Chinnappa, T. B., & Karunakaran, N. (2021). Customer relationship management and higher education in India. *Journal of Management Research and Analysis*, *8*(4), 180-188. https://doi.org/10.18231/j.jmra.2021.035

[14] Rigo, G., Pedron, C. D., Caldeira, M., & Araújo, C. C. S. (2016). CRM adoption in a higher education institution. *TECSI Information Systems and Technology Management*. University of Sao Paulo. https://doi.org/10.4301/S1807-17752016000100003

[15] Laudon, K., & Laudon, J. (2012). *Management Information Systems – Managing the Digital Firm*. (12th Ed.). Pearson Education Inc.

[16] Richards, K., & Jones, E. (2008). Customer relationship management: Finding value drivers. *Industrial Marketing Management*, *37*(2), 120-130. https://doi.org/10.1016/j.indmarman.2006.08.005

[17] Mendoza, L., Marius, A., Pérez, M., & Grimán, A. (2007). Critical success factors for a customer relationship management strategy. *Information and Software Technology*, *49*(8), 913-945. https://doi.org/10.1016/j.infsof.2006.10.003

[18] Reinartz, W., Krafft, M., & Hoyer, W. (2004). The customer relationship management process: Its measurement and impact on performance, *Journal of Marketing Research*, *41*(3), 293-305. https://doi.org/10.1509/jmkr.41.3.293.35991

[19] Oksanen, T. (2010). *CRM ja muutoksen tuska – Asiakkuudet haltuun*. Talentum Media Oy. Helsinki.

[20] Meyliana, Hidayanto, A. N., & Budiardjo, E. K. (2016). The critical success factors for customer relationship management implementation: a systematic literature review. *International Journal of Business Information Systems*, *23*(2), 131-174. https://doi.org/10.1504/IJBIS.2016.078904

[21] Ankrah, S., & Al-Tabbaa, O. (2015). Universities–industry collaboration: A systematic review, *Scandinavian Journal of Management*, 31(3), 387-408. https://doi.org/10.1016/j.scaman.2015.02.003

[22] Yin, R. K. (2014). *Case Study Research Design and Methods*. (5th ed.). Sage.

[23] Bryman, A., & Bell, E. (2011). *Business Research Methods*. (3rd ed.). Oxford University Press. Oxford

[24] Nguyen, B., & Mutum, D.S. (2012). A review of customer relationship management: successes, advances, pitfalls and futures, *Business Process Management Journal*, *18*(3), 400-419. https://doi.org/10.1108/14637151211232614

Authors' contacts:

Samuel Lakkala, M.Sc. (Tech)
(Corresponding author)
Department of Industrial Engineering and Management, University of Oulu,
PO Box 4610, FI-90014 Oulu, Finland
+358 40 526 1564, samuel.lakkala@oulu.fi

Joonas Suutari, M.Sc. (Tech)
joonas.suutari@live.fi

Dr Pekka Tervonen, Research director
Department of Industrial Engineering and Management, University of Oulu,
PO Box 4610, FI-90014 Oulu, Finland
pekka.tervonen@oulu.fi

Dr Harri Haapasalo, Professor
Department of Industrial Engineering and Management, University of Oulu,
PO Box 4610, FI-90014 Oulu, Finland
harri.haapasalo@oulu.fi

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240810013049

Original scientific paper
Received: 2024-08-10, Accepted: 2024-12-20

An Adaptive Load-Shedding Approach Applying Wide-Area Measurement Systems (WAMS)

Kazem Mehrabi, Mohammad Farhadi-Kangarlu*, Behrouz Tousi

Abstract: Power system collapses are among the primary challenges in modern grids. Today's power grids face stressful conditions day by day. The increase in electricity consumption, the expansion of electrical networks, and the oldness of protective equipment are among the stressful items in today's networks. In this stressful situation caused by the low stability margin in power networks, load shedding (LS) is the last solution to prevent power system collapse. Various conventional and independent LS programs will no longer respond to the existing conditions. Therefore, the load-shedding scheme considering the existing conditions and appropriate to the types of instabilities is very necessary. To improve the load-shedding programs, various methods are suggested in the articles. The expansion of wide-area measurement systems and increasing use in electrical networks have made it possible to be more maneuverable in load-shedding programs than before. Hence, a new adaptive LS approach is proposed based on the combined instabilities, which shows better behavior than the existing programs. Validation of the suggested approach is investigated in a real power grid.

Keywords: adaptive load shedding scheme; combined load shedding; adaptive protection system; wide-area measurement systems (WAMS)

1 INTRODUCTION

With the growing energy demand, power systems are now operating very close to their sustainability limits, leading to a minimal sustainability margin. As a result, network components are functioning near their maximum available capacity [1], [2]. In such circumstances, any significant disruption can cause an imbalance in consumption and production and ultimately jeopardize the power system's stability [3-5]. Therefore, in recent years, there have been numerous reports of various types of instabilities and even overall blackouts in power systems. Voltage instability, frequency instability, and rotor angle instability are the basis of power system instability. To assert such instabilities, several practical measures are implemented [6]. The first action involves activating stability control equipment to maintain the power grid's stability. Load-shedding (LS) is used as a last-resort action to preserve network stability [6].

Frequency is a key indicator in power systems, reflecting the balance between production and consumption of active power. One of the most widely used LS approaches to maintain power grid stability is the under-frequency load-shedding (UFLS) scheme. In this approach, the timing, location, and quantity of loads to be shed for each disturbance, regardless of its intensity, are determined by relying on previous network evaluations. Similarly, the under-voltage LS plan was developed to prevent voltage instability in power grids. Despite the local nature of voltage stability, this plan is also implemented based on offline evaluations [7].

Although conventional load-shedding plans can sometimes restore power systems' stability, they have several drawbacks. These include setting a frequency threshold, determining the shedded load amount, the stepping load-shedding plan, selecting loads leaning on previous network studies, and ignoring the combination of different inconsistencies, among other issues [8].

Multiple studies have introduced strategies to address the limitations of traditional load-shedding (LS) programs. Several researchers have suggested various schemes to minimize the drawbacks of the under-frequency load-shedding plan. These include the reduction of the relay's delay time during major disturbances with significant frequency drops and local LS are among other plans aimed at maintaining the power systems' frequency stability [9]. However, failing to account for the combined instabilities remains a significant defect of the designed plans. In references [10, 11], an LS program has been introduced, but the one-dimensional perspective on instabilities continues to be a major disadvantage of existing approaches. In [12], the authors applied the Thevenin equivalent model to enhance the conventional UFLS plan. However, one of the drawbacks of this approach is its high computational burden. References [13, 14] suggested utilizing voltage changes to precisely determine the locations for load shedding within the voltage load-shedding program. The main defect of many suggested cases is that they rely on the offline evaluation of loads leading to instability, take gradual load increase as the agent of instability, and do not consider the combined instabilities. To accurately determine the active power unbalance, response-based and event-based approaches have been introduced. In reference [15], the voltage stability index (VSI) is utilized to recognize buses as vulnerable to instability. This local indicator is effective for assessing voltage instability but might not be suitable for addressing combined voltage and frequency instabilities. Simultaneously, determining the networks' impedance and generator equivalents to determine the location of shed loads using the Jacobian Matrix is another method that can be applied in power grids. However, the high computational time makes its online application challenging [16]. Seyedi et al. introduced load-shedding methods based on responses and events to determine the amount of load to be removed. In this approach, they proposed the LS plan in specific steps, tailored for both large and small events [17]. Even though computational time has been reduced and stability has relatively improved, the proposed method still faces challenges such as voltage amplitude index and the presence of buses with identical magnitudes.

Voltage instability often does not occur in isolation; it is

frequently intertwined with angel instability, where one can be unstable the other. Similarly, frequency instability may also be highly dependent on either angle or voltage instabilities. Both voltage and frequency instability are critical factors in power system instability, and if not addressed promptly, they can cause other power plant units to trip, potentially leading to a partial or complete blackout of the power grid.

With the growing adoption of WAMS in power grids, control and protection programs have gained greater flexibility. Therefore, in this manuscript, a new adaptive and wide-area load-shedding scheme is proposed that takes into account the combination of various instabilities. This method also considers the impact of rotor angle on instabilities.

The rest of the manuscript structure is as follows: Section 2 introduces both adaptive and conventional UFLS schemes. Section 3 provides a detailed evaluation of the proposed LS scheme. Section 4 briefly introduces and analyzes the studied test system along with the related figures. Finally, Section 5 presents the conclusions.

2 LOAD SHEDDING (LS) APPROCHES

Various LS approaches have been suggested to enhance LS schemes. The conventional UFLS method is the oldest LS plan in power grids, while the modal analysis plan, highlighted in many recent articles, is presented as a new approach and thoroughly reviewed.

2.1 Conventional Under- Frequency LS Method

In this LS approach, when the frequency falls under a specified threshold, it will sequentially disconnect certain loads from specific buses. The overall power imbalance will be calculated using Eq. (1) and Eq. (2).

$$L = \frac{Total\ Consumption\ Load - Total\ Generation\ Power}{Total\ Generation\ Power} \quad (1)$$

$$LD = \frac{\frac{L}{L+1} - d\left[1 - \left(\frac{f_{\min}}{f_{\mathrm{n}}}\right)\right]}{1 - d\left[1 - \left(\frac{f_{\min}}{f_{\mathrm{n}}}\right)\right]} \quad (2)$$

Here, D and LD illustrate the load dependence on frequency changes and the amount of shedded load in specific steps [17].

2.2 Modal Analysis Based Load Shedding Method

To pinpoint the precise location of loads inclined to instability, the Modal analysis approach was suggested in [14]. The active and reactive power (load) is determined using (3). In the following equation, J represents the Jacobian matrix.

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = \begin{bmatrix} J_{P\theta} & J_{PV} \\ J_{Q\theta} & J_{QV} \end{bmatrix} \begin{bmatrix} \Delta\theta \\ \Delta V \end{bmatrix} \quad (3)$$

Taking into account the significant affiliation of voltage variations with reactive power variation, and considering certain hypotheses, the following equations are obtained.

$$\Delta Q = \left[J_{QV} - J_{Q\theta} J_{P\theta}^{-1} J_{PV} \right] \Delta V = J_{\mathrm{R}} \Delta V \quad (4)$$

$$\Delta V = J_{\mathrm{R}}^{-1} \Delta Q \quad (5)$$

In this equation, J_{R} is the reduced Jacobian matrix. Based on Eq. (6) and by breaking it into left and right components, Eq. (7) and Eq. (8) are derived.

$$J_{\mathrm{R}}^{-1} = E_{\mathrm{R}} \xi^{-1} E_{\mathrm{L}} \quad (6)$$

$$\Delta V = E_{\mathrm{R}} \xi^{-1} E_{\mathrm{L}} \Delta Q = \sum_{l} \frac{E_{\mathrm{R},\,l} E_{\mathrm{L},\,l}}{\lambda_l} \Delta Q \quad (7)$$

$$\Delta V_{ml} = \frac{\Delta Q_{ml}}{\lambda_l} \quad (8)$$

In the above equations, E_L and E_R represent the left and right eigenvectors of J_R and also $E_{\mathrm{L},l}$ and $E_{\mathrm{R},l}$ are indicators of the l^{th} row and column of E_{R}.

Here, λ_l being less than or equal to zero indicates the presence of voltage instability. By decomposing k to the related eigenvectors, the voltage variation due to reactive power variation is represented by Eq. (9).

$$VQS_k = \frac{dV_k}{dQ_k} = \sum_{l} \frac{\mu_{k,l} \cdot \eta_{l,k}}{\lambda_l} \quad (9)$$

Each bus that VQS_k (V-Q sensitivity at bus k) amounts to lean to zero, is more prone to voltage instability.

This approach is reliable. However, it has some disadvantages. For example:

- The high computational burden of the Modal analysis makes it difficult to use this method on online platforms, especially in large or combined disturbances.
- The agent of instability is limited to gradual load growth, while generator trips and line outages have an important role in power system instabilities.

3 THE NEW ADAPTIVE AND WIDE-AREA LS SCHEME

In [18] and [19], two types of methodologies are utilized to precisely quantify the power imbalance. The author proposed two-step thresholds of the frequency to configure the LS plan. However, the voltage magnitude index is proposed to determine the location of the shed loads. While this index can be applied online during disturbances, it is not a reliable measure for identifying buses prone to instability. Often, multiple buses have the same voltage magnitude, complicating the selection process for identifying buses with more instability potential.

In modern grids, the spreading use of WAMS increases the maneuverability of power networks. Hence, this manuscript introduces a combined scheme that considers the

frequency, voltage, and angle of various buses within the power system. Generators' rotor angle variations affect the voltage angle of neighboring buses and can aid the LS scheme. In the newly suggested approach, the exact power imbalance is calculated using wide area measurement systems data and based on the response-based method. As in [17], a two-step adjustment is used to determine the LS frequency threshold. In a network with a nominal frequency of 50 Hz, this threshold is designed based on the rate of change of frequency (ROCOF) in two steps of 49.7 and 49.4 Hz for large and small disturbances. The process of the suggested LS approach is shown in Fig. 1.

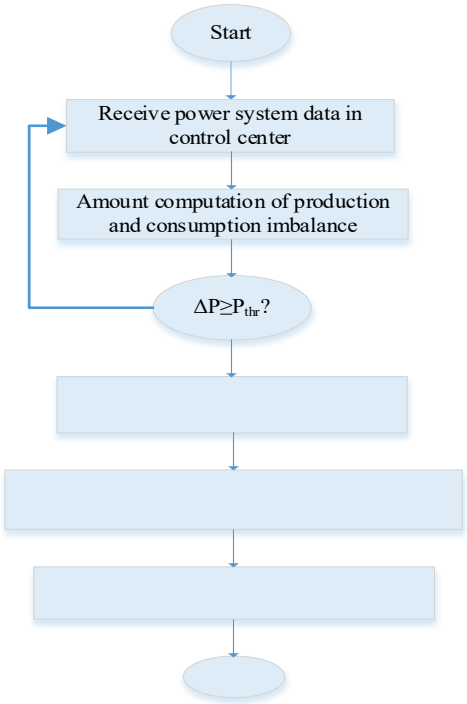


Figure 1 The process of the suggested LS approach

In the suggested approach, power imbalances based on power swing equations are calculated as follows. Here, P_{m_i} and P_{e_i} represent the generator's mechanical and electrical active power on i^{th} the bus.

$$\frac{2H_i}{f_n}\frac{df_i}{dt} = P_{m_i} - P_{e_i} = \Delta P_i \quad (i = 1, ..., N) \tag{10}$$

To account for the power imbalance across all generators, Eq. (11) can be derived.

$$\Delta P = \sum_{i=1}^{N} \Delta P_i = \frac{2\sum_{i=1}^{N} H_i}{f_n}\frac{df_c}{dt} = \varepsilon \frac{df_c}{dt} \tag{11}$$

Where

$$f_c = \frac{\sum_{i=1}^{N} H_i f_i}{\sum_{i=1}^{N} H_i} \tag{12}$$

$$\varepsilon = \frac{2}{f_n}\sum_{i=1}^{N} H_i \tag{13}$$

Considering the reserve capacity (P_{thr}) of the power plant units, the final load-shedding amount (P_{shed}) is determined using Eq. (15).

$$P_{shed} = SM \times (\Delta P - P_{thr}) \tag{14}$$

The tripped units may contribute to the spinning reserve included in the LS program. Therefore, a safety margin (SM) has been weighed into the LS calculations.

Given the localized essence of voltage and the presence of adjacent buses within the similarity limit, selecting buses to shed becomes challenging. To address this issue, a newly adaptive and wide-area LS scheme has been proposed. Fig. 2 illustrates the perspective of an equivalent grid. The voltage magnitude and angle of the bus are key factors in determining the current flow, which subsequently affects the voltage decrease on the power system. Consequently, as reactive and active power flows through the equivalent network, the voltage drop will be intensified, potentially leading to instability and collapse of the power system.

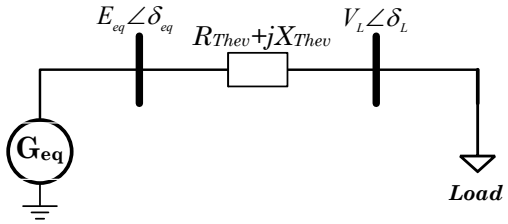


Figure 2 The perspective of an equivalent grid

The load current flow in the equivalent grid (I_L) is derived using Eq. (15).

$$I_L = \frac{E_{eq}\angle\delta_{eq} - V_L\angle\delta_L}{R_{Thev} + jX_{Thev}} \tag{15}$$

$$S_{GL} = E_{eq} I_L^* \tag{16}$$

In the above equations, E_{eq}, R_{Thev} and X_{Thev} illustrate the Thevenin equivalent of power network voltage, resistance and reactance, respectively, and S_{GL} represents the total power transmitted to the load.

Neglecting the line resistance in Eq. (15), the reactive and active power flow to the load from the equivalent network generator (P_{GL} and Q_{GL}) can be expressed as follows.

$$P_{GL} = \frac{E_{eq}V_L}{X_{Thev}}\sin(\delta_{eq} - \delta_L) \quad (17)$$

$$Q_{GL} = \frac{E_{eq}^2}{X_{Thev}} - \frac{E_{eq}V_L}{X_{Thev}}\cos(\delta_{eq} - \delta_L) \quad (18)$$

Because voltage is a local parameter, different buses in the specific area will have almost the same magnitude. Hence, a bus with a higher voltage angle will drag more current through the equivalent grid, leading to increased reactive and active power transfer through the equivalent network, resulting in a greater voltage drop. Simply put, for buses that are of a similar voltage magnitude, if the equivalent network voltage is supposed to be $1\angle 0$, the near the bus voltage angle is 90°, the higher the flow of reactive and active power and the result is the greater voltage falls. Consequently, that bus can unstable the power grid. The same principle applies to generators; the rotor angle in the vicinity of 90°, the more susceptible to outage.

Furthermore, to enhance the validity of the LS scheme, the suggested approach offers an adaptive and wide-area LS with a lower computation burden and a swift mechanism. This adaptive LS approach has improved the voltage and frequency stability of the power system while addressing the shortcomings of previous programs.

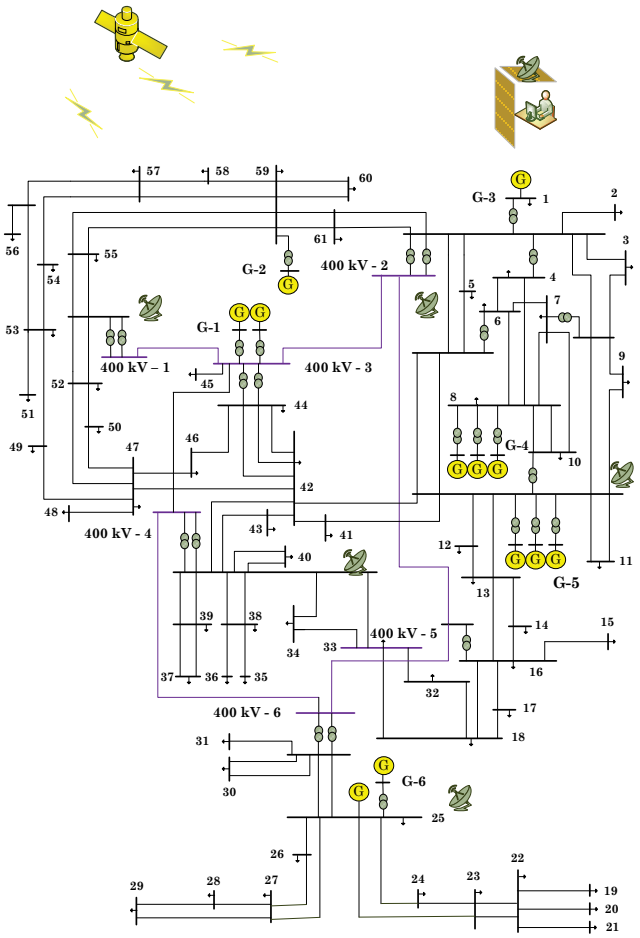


Figure 3 The real Khorasan power grid.

4 NUMERICAL EVALUATION AND DISCUSSIONS

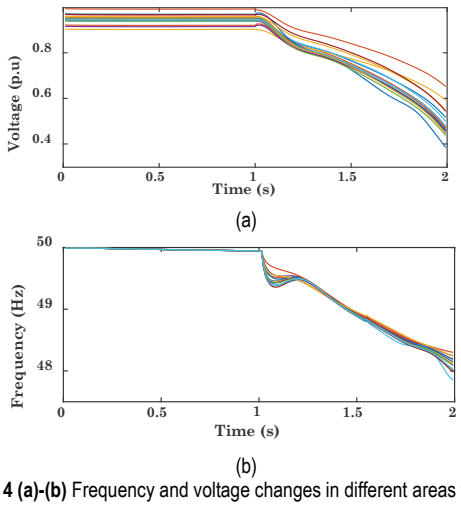
To appraise the effectiveness of the proposed LS approach, the Khorasan power grid (a major part of the Iranian national network) is analyzed using the DigSILENT platform. Fig. 3 depicts the simulated power grid.

This network with 75 buses, a generation capacity of 2370 MW and a spinning reserve of 350 MW, has been evaluated under different scenarios to evaluate the performance of the proposed LS scheme compared to existing approaches. In conventional UFLS relays, the intentional delay time is set at 100 ms, while the implementation time for the proposed scheme and the modal analysis method (assuming real-time implementation without the drawbacks of static analysis and computational load) is set at 200 ms.

The article presents two example scenarios with different characteristics and methods, which can be evaluated in the subsequent sections.

4.1 Major Event: Tripping of all Power Plant Units in G-1

In this disturbance, which occurred on the 1st second, a power of 906 MW was outage from the production capacity. Figs. 4 (a)-(b) illustrate the frequency and voltage changes in different zones without considering the LS approaches. As observed, notwithstanding the grid's spinning capacity and remedial actions, the system experiences concurrent frequency and voltage instabilities, potentially leading to power system collapse. Therefore, to prevent power system collapse, an appropriate load-shedding (LS) scheme must be implemented.

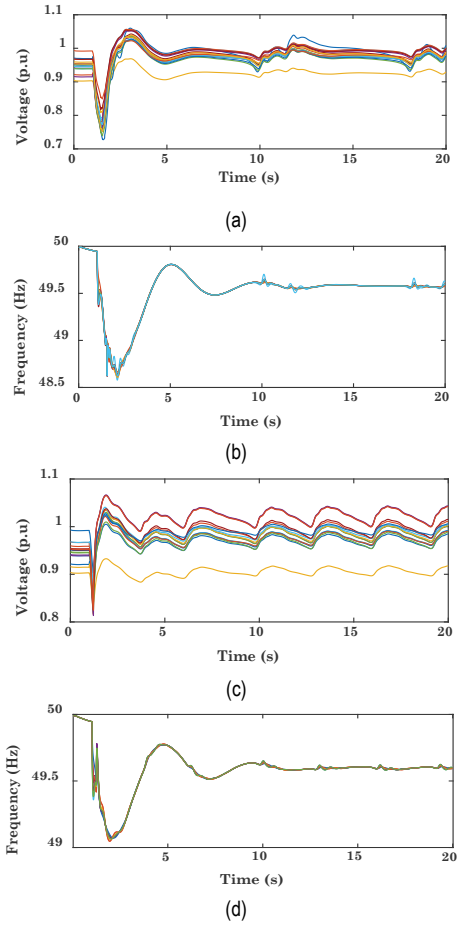


Figures 4 (a)-(b) Frequency and voltage changes in different areas following the major event.

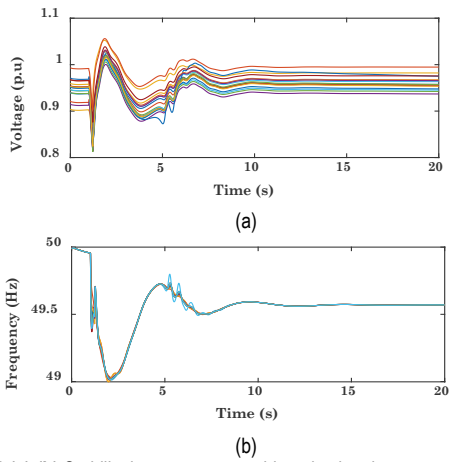
All the LS strategies analyzed, including conventional UFLS, the Modal analysis method, and the suggested approach, have been implemented on the power grid. Figs. 5 (a)-(d) depict the frequency and voltage sample signals in various areas of the grid for conventional UFLS and Modal analysis-based LS approaches. Observing Fig. 5 closely, it becomes evident that although the two simulated load-

shedding methods can restore voltage and frequency stability, there are persistent oscillations in the frequency and voltage waveforms. This indicates that the shed loads might not be optimally distributed.

The frequency and voltage signals for different areas of the power system under the proposed LS approach are depicted in Figs. 6 (a)-(b). Implementing this approach stabilizes the frequency and voltage signals, as evidenced by the absence of persistent state fluctuations.

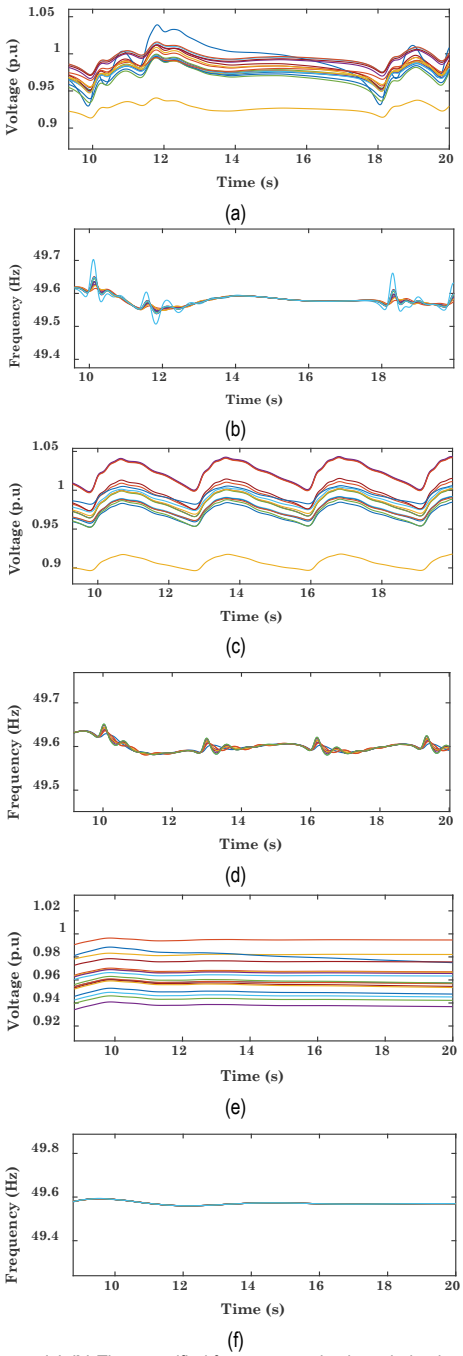


Figures 5 (a)-(b) Stability improvements achieved using the conventional UFLS scheme, **(c)-(d)** Stability improvements achieved using the Modal analysis scheme.



Figures 6 (a)-(b) Stability improvements achieved using the proposed LS scheme.

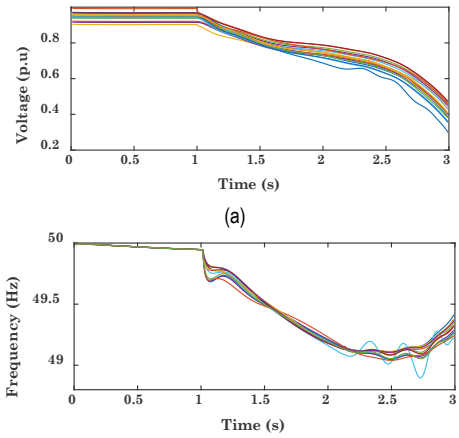
Based on the simulation results and the magnification of the frequency and voltage signals for sample buses in various areas during a near-to-final condition (from the 10th to the 20th second), Figs. 7 (a)-(f) have been derived. Upon close examination, it becomes evident that, notwithstanding their ability to restore frequency and voltage, the conventional UFLS and modal analysis methods also exhibit oscillatory instabilities. In contrast, the proposed scheme achieves better frequency and voltage recovery with less amplitude variation and no oscillation.



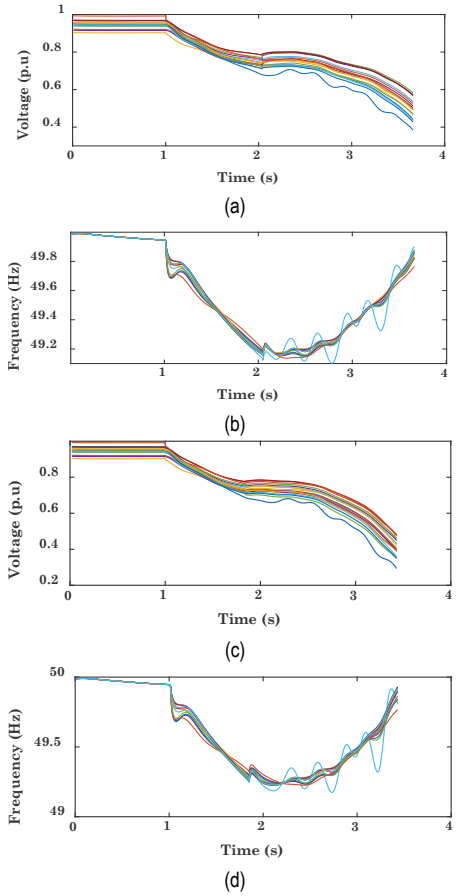
Figures 7 (a)-(b) The magnified frequency and voltage behavior for the conventional UFLS approach, **(c)-(d)** The magnified frequency and voltage behavior based on the Modal analysis LS approach, **(e)-(f)** The magnified frequency and voltage behavior based on the proposed LS approach.

4.2 Minor Event: Tripping of all Power Plant Units in G-3 and 2 to 3 Tie-Line

In this disturbance, which occurred on the 1st second, a power of 560 MW was outage from the production capacity. Subsequently, the 400 kV-2 to 3 tie-line is also outage. Figs. 8 (a)-(b) display the frequency and voltage variations across different areas of the network.

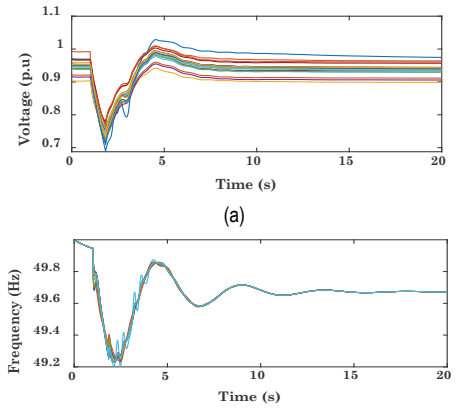


Figures 8 (a)-(b) Frequency and voltage changes in different areas following the minor event.



Figures 9 (a)-(b) Stability improvements achieved using the conventional UFLS scheme, **(c)-(d)** Stability improvements achieved using the Modal analysis plan.

The frequency and voltage outcomes in various areas of the power system, derived from LS approaches, are illustrated in Figs. 9 (a)-(d) and Figs. 10 (a)-(b).



Figures 10 (a)-(b) Stability improvements achieved using the proposed LS scheme.

As depicted in the corresponding figures, both the conventional UFLS and Modal analysis approaches, despite their efforts to recover frequency, have experienced voltage instability, resulting in the network's collapse. The proposed index demonstrates superior performance among the simulated methods.

4.3 General Discussions

A detailed brief of the simulation results is provided in Tab. 1. For the evaluated network, the threshold value used to identify the severity of large or small events is based on the ROCOF, set at 1.7 Hz/s and 1.4 Hz/s, respectively. In this simulated system, a safety margin is considered at 1.05. Events in this simulation are initiated at the 1st second. If stability is observed around the 20th second, Tab. 1 will display the voltage and frequency magnitude range (V_{min}, V_{max} and F_{min}, F_{max}) of the network's various buses following the different events. In the first event, the conventional load-shedding plan is executed in three steps, resulting in higher shed loads compared to other approaches. However, in the absence of adaptive selection for the shed load locations, voltage and frequency waveforms exhibit fluctuations in various areas during steady states. Similarly, despite the computational burden of the Modal analysis LS approach, it also results in voltage and frequency fluctuations, although with fewer shed loads than the conventional method.

The proposed LS scheme, on the other hand, demonstrates superior and acceptable behavior post-disturbance. With fewer shed loads, it achieves better frequency and voltage restoration across different network areas, exhibiting low variance and no fluctuations.

In the other scenario, the conventional LS approach is activated only in the first step, leading to voltage instability and eventual network collapse despite efforts to restore frequency. The modal analysis plan shows similar behavior, causing network instability and collapse due to voltage issues despite shedding more loads.

The proposed load-shedding scheme, with lower computational demands compared to the modal analysis and more adaptive to the type and location of the incident than conventional methods, performs acceptably in such combined disturbance scenarios. In both modal analysis and conventional UFLS approaches, the network is prone to instability and blackouts.

The proposed LS approach is feasible for implementation in any power network, thanks to the availability of WAMS. Future research can explore the impact of load models and load restoration methods on the load-shedding program following network modifications.

Table 1 Comparison of the simulated LS approaches

Disturbance	LS approach	F_{min} (Hz)	F_{max} (Hz)	V_{min} (p.u)	V_{max} (p.u)	ΔP_{shed} (MW)	Shedded buses
First scenario	Conventional	49.576	49.625	0.929	1.008	542.9	4-11-14-15-28-36-37-40-43-47-52-53-54-55
	Modal analysis	49.463	49.599	0.903	1.019	583.8	1-2-3-6-9-12-15-17-18-20-21-22-23-25-27-32-34-42-44-50-51-53-54-55-58-59-60
	Proposed	49.569	49.569	0.937	0.995	531.1	6-7-10-37-48-49-51-53-54
Second scenario	Conventional	Unstable	Unstable	Unstable	Unstable	158.7	15-40-47-53-54
	Modal analysis	Unstable	Unstable	Unstable	Unstable	221.5	3-35-46
	Proposed	49.671	49.671	0.899	0.975	219.4	2-16-17-19-20-21-22

5 CONCLUSION

In summary, this study proposes a novel adaptive LS approach designed to enhance power system frequency and voltage stability. Key insights include:

- **WAMS Data**: Fundamental for implementing the scheme, facilitating precise selection of loads prone to instability.
- **Reduced Computational Burden**: By considering another network index for load selection, the method reduces computational demands compared to other adaptive programs.
- **Improved Response**: The adaptive approach effectively addresses various types of disturbances, outperforming constant load-shedding patterns used in conventional methods.
- **Advantages over Conventional UFLS**: Avoids the limitations of fixed load-shedding amounts, timing, and locations.
- **Challenges with Modal Analysis**: Although providing an admitted result, it entails a high computational burden and may not always maintain network stability, because of gradual load growth as an instability factor.
- **Better Performance**: The proposed scheme shows faster and superior behavior in responding to disturbances, aiming to mitigate the shortcomings of existing programs.

The adaptive load-shedding scheme leverages WAMS data and a strategic load selection index to offer a more resilient and efficient solution for maintaining power system stability.

5.1 Notice

The Khorasan network, encompassing 5 large power plants and 75 buses at 400 kV and 132 kV levels, boasts an installed generation capacity of approximately 2700 MW, with a spinning reserve of around 350 MW at peak load. For dynamic network stability studies in this paper, the Khorasan network has been meticulously and dynamically modeled, incorporating an accurate load model that simulates load behavior under voltage and frequency instability. Additional details about the Khorasan network are available in reference [20].

Nomenclature

L	Total overload amount
LD	Total amount of load to be shed in UFLS
d	Load reduction factor
f, $\mathrm{d}f/\mathrm{d}t$	Frequency and its rate of change (ROCOF)
f_{min}	Minimum permissible frequency in power network
f_n	Power system nominal frequency
f_c	Frequency of the equivalent inertial centre
$f_{setting}$	Frequency setting in UFLS scheme
N	Total number of generators
H_i	Inertia constant in i^{th} generator
V	Bus voltage magnitude
i, k	Bus indices

Greek Symbols

ε	Constant value
ΔP	Generation-consumption mismatch
δ_L	Bus voltage angle
ξ	Diagonal eigenvalue of JR
ΔV_{ml}	Variation of l^{th} modal voltage
ΔQ_{ml}	Variation of l^{th} modal reactive power
λ_l	l^{th} eigenvalue of J_R
$\mu_{k,l}$	k^{th} element of $E_{R,l}$
$\eta_{k,l}$	k^{th} element of $E_{L,l}$

6 REFERENCES

[1] Zheng, M., & Yu, J. (2024). Utilization of "Intersection" Methodology for Optimization with Many Objectives in Designed Experiments of Material Processing. *Tehnički glasnik*, *18*(3), 337-341. https://doi.org/10.31803/tg-20220426135403

[2] Shekari, T., Aminifar, F., & Sanaye-Pasand, M. (2015). An analytical adaptive load shedding scheme against severe combinational disturbances. *IEEE Transactions on Power Systems, 31*(5), 4135-4143. https://doi.org/10.1109/TPWRS.2015.2503563

[3] Kampo, M., Ndiaye, B. M., & Degla, G. (2018). Dynamic Programming for an Optimal and Equitable Public Load Shedding Schedule. *Croatian Operational Research Review, 9*(2), 211-221. https://doi.org/10.17535/crorr.2018.0016

[4] Toro-Mendoza, M. A., Segundo-Ramírez, J., Esparza-Gurrola, A., Visairo-Cruz, N., Guitiérrez, C. A. N., & Pérez-Negrón, C. (2023). Toward Adaptive Load Shedding Remedial Action Schemes in Modern Electrical Power Systems. *IEEE Access*, 11, 111011-111033. https://doi.org/10.1109/ACCESS.2023.3322657

[5] Chen, R., Xu, H., Zhou, L., Cai, J., Xiong, C., Zhou, Y., ..., & Yang, N. (2023). Underfrequency Load Shedding Strategy With an Adaptive Variation Capability for Multi-Microgrids. *IEEE Access, 11*, 17294-17304. https://doi.org/10.1109/ACCESS.2023.3246088

[6] Chandra, A., & Pradhan, A. K. (2020). An adaptive underfrequency load shedding scheme in the presence of solar photovoltaic plants. *IEEE Systems Journal, 15*(1), 1235-1244. https://doi.org/10.1109/JSYST.2020.2995050

[7] Dedović, M. M., Avdaković, S., Musić, M., & Kuzle, I. (2024). Enhancing power system stability with adaptive under frequency load shedding using synchrophasor measurements and empirical mode decomposition. *International Journal of Electrical Power & Energy Systems*, 160, 110133. https://doi.org/10.1016/j.ijepes.2024.11013

[8] Abrar, S. F., Masood, N. A., & Alam, M. J. (2024). An adaptive load shedding methodology for renewable integrated power systems. *Heliyon, 10*(21). https://doi.org/10.1016/j.heliyon.2024.e40043

[9] Abedini, M., Sanaye-Pasand, M., & Azizi, S. (2014). Adaptive load shedding scheme to preserve the power system stability following large disturbances. *IET Generation, Transmission & Distribution*, 8(12), 2124-2133. https://doi.org/10.1049/iet-gtd.2013.0937

[10] Rather, Z. H., & Pal, B. C. (2023). Online estimation of disturbance size and frequency nadir prediction in renewable energy integrated power systems. *IEEE Transactions on Power Systems*, 39(1), 1126-1137. https://doi.org/10.1109/TPWRS.2023.3267744

[11] Schneider, K. P., Sun, X., & Tuffner, F. (2023). Adaptive load shedding as part of primary frequency response to support networked microgrid operations. *IEEE Transactions on Power Systems, 39*(1), 287-298. https://doi.org/10.1109/TPWRS.2023.3261222

[12] Luo, Z., Liu, H., Wang, N., Zhao, T., & Tian, J. (2024). Optimal adaptive decentralized under-frequency load shedding for islanded smart distribution network considering wind power uncertainty. *Applied Energy*, 365, 123162. https://doi.org/10.1016/j.apenergy.2024.123162

[13] He, P., Wen, B. & Wang, H. (2019). Decentralized adaptive under frequency load shedding scheme based on load information. *IEEE Access, 7*, 52007-52014. https://doi.org/10.1109/ACCESS.2019.2911665

[14] Jalilzadeh, S., hadi Hosseini, S., & Derafshian-Maram, M. (2010, October). Optimal load shedding to prevent voltage instability based on multi-objective optimization using modal analysis and PSO. In *IEEE International Congress on Ultra Modern Telecommunications and Control Systems*, 371-376. https://doi.org/10.1109/ICUMT.2010.5676611

[15] Rudez, U., & Mihalic, R. (2010). Monitoring the first frequency derivative to improve adaptive underfrequency load-shedding schemes. *IEEE Transactions on Power Systems, 26*(2), 839-846. https://doi.org/10.1109/TPWRS.2010.2059715

[16] Laghari, J. A., Mokhlis, H., Karimi, M., Bakar, A. H. A., & Mohamad, H. (2014). A new under-frequency load shedding technique based on combination of fixed and random priority of loads for smart grid applications. *IEEE Transactions on Power Systems, 30*(5), 2507-2515. https://doi.org/10.1109/TPWRS.2014.2360520

[17] Pasand, M. S., & Seyedi, H. (2007, October). New centralized adaptive under frequency load shedding algorithms. In *IEEE Large Engineering Systems Conference on Power Engineering*, 44-48. https://doi.org/10.1109/LESCPE.2007.4437350

[18] Mehrabi, K., Golshannavaz, S., & Afsharnia, S. (2019). An improved adaptive wide-area load shedding scheme for voltage and frequency stability of power systems. *Energy Systems, 10*, 821-842. https://doi.org/10.1007/s12667-018-0293-9

[19] Li, Q., Xu, Y., Ren, C., & Zhang, R. (2022). A probabilistic data-driven method for response-based load shedding against fault-induced delayed voltage recovery in power systems. *IEEE Transactions on Power Systems, 38*(4), 3491-3503. https://doi.org/10.1109/TPWRS.2022.3206839

[20] Abedini, M. (2013). Combination of Voltage and Frequency Signals for Adaptive Load Shedding in Power Grids. *Master's Thesis, University of Tehran*.

Authors contacts:

Kazem Mehrabi
Faculty of Electrical and Computer Engineering, Urmia University, Urmia, Iran
E-mail: k.mehrabi@urmia.ac.ir

Mohammad Farhadi-Kangarlu
(Corresponding author)
Faculty of Electrical and Computer Engineering, Urmia University, Urmia, Iran
E-mail: m.farhadi@urmia.ac.ir

Behrouz Tousi
Faculty of Electrical and Computer Engineering, Urmia University, Urmia, Iran
E-mail: b.tousi@urmia.ac.ir

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250315172439

Original scientific paper
Received: 2025-03-15, Accepted: 2025-05-20

Sustainable Supply Chain Management, Information Technology, and Firm Performance: The Mediating Power of Agility and Supply Chain Integration

Asiye Omidian, Athena Omidian*, Shahriar Mohajer, Elnaz Zarei, Azita Ramezani

Abstract: This study investigates how sustainable supply chain management (SCM) practices and information technology (IT) drive firm performance in small and medium enterprises (SMEs), with a focus on the mediating roles of supply chain agility (SCA) and sustainable supply chain integration (SSCI). Drawing on data from 219 managers and supply chain experts in Iranian SMEs, structural equation modeling (SEM) revealed that both sustainable SCM practices and IT adoption significantly enhance firm performance directly and indirectly through SCA and SSCI. Sustainable SCM practices demonstrated a robust influence on SSCI ($\beta = 0.35$, $p < 0.001$), SCA ($\beta = 0.49$, $p < 0.001$), and firm performance ($\beta = 0.44$, $p < 0.001$), while IT similarly impacted SSCI ($\beta = 0.34$, $p < 0.001$), SCA ($\beta = 0.39$, $p < 0.001$), and performance ($\beta = 0.21$, $p < 0.001$). Critically, SCA and SSCI emerged as pivotal mediators, explaining 48% of the variance in firm performance. The findings underscore that agility and integration transform sustainable and technological investments into competitive outcomes, enabling SMEs to navigate dynamic markets. Practical insights highlight the need for firms to prioritize cross-departmental coordination, IT-driven customer insights, and adaptive decision-making frameworks.

Keywords: firm performance; information technology; supply chain agility; supply chain integration; supply chain management; sustainable

1 INTRODUCTION

The concept of a firm's performance is quite extensive, including not only the results produced by the organization but also various other factors that contribute to its overall effectiveness. Alternatively, a firm's performance refers to its ability to accomplish its mission and activities and to achieve its objectives [1]. In competitive markets, organizational leadership consistently prioritizes performance enhancement as a strategic necessity. This imperative stems from the understanding that sustained operational refinement is vital for maintaining relevance—not merely to survive but to thrive amid shifting industry demands and rival market forces. Therefore, Adapting and implementing effective strategies and continuously improving performance is vital for companies in today's changing, complex, and uncertain competitive environment). Factors such as accelerated environmental shifts, fiscal constraints, organizational streamlining processes, and heightened public demands for more accountability of organizations for their performance, a greater emphasis has been placed on performance management as a result. Due to various factors including swift changes in the environment, budget shortfalls, organizational downsizing and restructuring, as well as increasing societal demands for accountability in performance, there has been a heightened focus on performance management. In other words, one of the main topics in the organizational analysis is performance and there is a constant search for ways to improve the performance of companies [2]. Therefore, the central goal of this paper is to explore the influence that sustainable SCM practices and information technology (IT) may have on performance of firm by taking into account the mediation role that agility and sustainable supply chain integration (SSCI) may play in determining firm performance.

SCM integrates supply chain activities to continuously achieve competitive advantage as well as their information by improving chain relationships. Therefore, SCM is a process that integrates supply chain activities and related information currents by improving and coordinating activities in a chain of production and supply. Adding sustainability to SCM has expanded its concept [3]. Sustainability is defined as a process by which a balance is maintained between meeting current needs and meeting future generations' needs. A sustainable SCM manages material, information, and capital flow and coordinates organisations as part of the supply chain, and integrates underlying principles of sustainable development that relate to the economy, the environment, and society, resulting from the needs of customers and beneficiaries [4]. To ensure long-term viability in supply chain operations, participants in sustainable supply chains must balance dual imperatives: adherence to rigorous social and environmental standards, and the preservation of economic competitiveness. This necessitates aligning with consumer expectations while fulfilling key economic benchmarks to sustain market relevance [5].

Sustainable SCM has emerged as a new and important approach by using environmental risk reduction to gain market share and profit. Organizations that intend to implement sustainable SCM need to review different strategies to reduce energy costs [6]. A sustainable SCM controls environmental effects in its product life cycle by emphasizing the integration of environmental management and SCM coordinates supply chain members, and encourages information sharing between them [7]. The performance of firms can be significantly impacted by sustainable SCM, according to studies [8].

Currently, IT is an incentive for the globalization of products and markets, and it increases the dynamics of services and monetary-financial currents, and often allows the increasing performance and efficiency [9]. A glance at the developed countries reveals that IT has played a vital role in their comprehensive development. Comprehensive

information and IT used at the right time prevents resource waste and unnecessary and unimportant tasks and most importantly, ineffective and inefficient decisions. In addition, technology is one of the most valuable resources of organizational processes to improve growth and productivity. Therefore, Improved performance through better allocation of technology resources has emerged as a significant subject in contemporary research literature. Overall, a significant amount of literature has been written on the importance of IT-related skills, competencies, and capabilities due to the important role they play in the performance of companies. Evidence suggests that IT has a positive influence on firm performance [10]. IT adoption not only optimizes operational efficiency but also supports customer-centric strategies, such as participatory digital marketing models facilitated by knowledge management [24, 25], which can further amplify supply chain responsiveness.

A supply chain's agility (SCA) is becoming increasingly important in response to short-term changes in the supply chain as a result of these changes; in fact, agility means that in order to adapt quickly, the supply chain has to be flexible to changes and can respond appropriately to them [11]. In other words, agility enables the supply chain to face unexpected changes. Companies realize that they need agility to survive and remain competitive. In addition, not every company has available resources to deal with different opportunities. Therefore, global competition requires organizations to Provide support to suppliers and customers by aligning their operations to achieve higher agility than other organizations; this integration is called "agile supply chain" [12]. SCA promotes the adaptability and flexibility of businesses and enables them to respond to market changes quickly and effectively. As a winning strategy, advocates of SCA recommend businesses use it as their supply model to compete at national and international levels [13]. Results of the conducted research have also shown that the performance of firms is positively affected by SCA [14, 15].

By integrating the supply chain process, all barriers tend to be removed in order to enable the smooth flow of materials, resources, and information throughout the supply chain. In order to reduce ambiguity in terms of integration, sharing information contributes to improving the capability to predict and reduce costs [16]. Supply chain integration (SCI) emphasizes high levels of customer-supplier engagement to reduce product development costs through the involvement of primary suppliers [17]. Results of the conducted studies have also shown that an important factor in a firm's performance can be attributed to SCI [18].

Theoretical research consistently underscores the significant impact of Sustainable SCM practices, IT adoption, operational agility, and SSCI on organizational performance. However, empirical studies reveal a critical gap: limited evidence has established a comprehensive framework linking sustainable SCM practices and IT capabilities to performance outcomes, particularly through the mediating effects of agility and SSCI. To bridge this gap, constructing a holistic model is imperative. Such a framework would elucidate the pathways through which IT investments enhance firm performance, with agility and SSCI serving as pivotal intermediaries in this relationship.

2 CONCEPTUAL MODEL

A conceptual framework (Fig. 1) was derived through an analysis of existing literature. This model positions Sustainable SCM practices and IT as independent variables, while SCA and SSCI function as mediating factors. Firm performance is established as the dependent variable, capturing the cumulative impact of these interrelated components. Therefore, these hypotheses are presented in the following order:

H1: Sustainable SCM practices are effective on SSCI.

H2: Sustainable SCM practices are effective on SCA.

H3: Sustainable SCM practices are effective on firm performance.

H4: IT is effective on SSCI.

H5: IT is effective on SCA.

H6: IT is effective on firm performance.

H7: SSCI is effective on SCA.

H8: SSCI is effective on firm performance.

H9: SCA is effective on firm performance.

H10: SSCI and SCA mediate the influence of sustainable SCM practices on firm performance.

H11: SSCI and SCA mediate the influence of IT on firm performance.

H12: The effect of SSCI on firm performance is mediated by SCA.

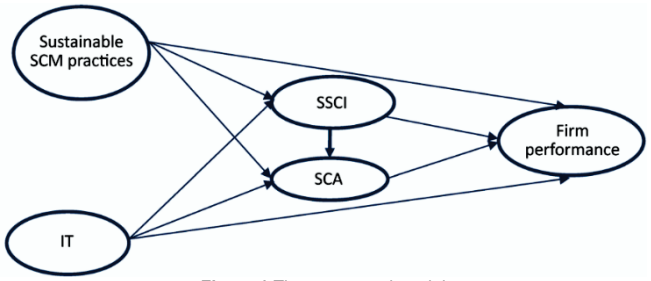


Figure 1 The conceptual model

3 RESEARCH METHODOLOGY

This study employed a descriptive correlational design combined with structural equation modeling (SEM) to analyze the hypothesized relationships between variables. SEM was selected for its capacity to evaluate complex, multivariate dependencies within the proposed framework [20, 21, 27].

3.1 Population and Sample

The target population comprised managerial personnel, deputies, and supply chain specialists within small-to-medium enterprises (SMEs) operating in Iran. A total of 250 survey questionnaires were distributed to eligible participants in these roles. After excluding incomplete responses, 219 valid questionnaires were retained for analysis. Responses were collected using a five-point Likert

scale, with anchors ranging from 1 ("strongly disagree") to 5 ("strongly agree").

3.2 Instruments

Sustainable SCM practices: Sustainable SCM practices were evaluated using a validated instrument adapted from Hong et al. [8]. The tool comprises 16 items grouped into five dimensions: trust and coordination within the supply chain (3 items), collaborative learning across supply chain partners (3 items), strategic orientation of supply chain operations (3 items), proactive risk management in supply networks (4 items), and continuity planning for supply chain resilience (3 items).

IT: Jiang and McCabe's [19] questionnaire was used to measure IT in 6 items; 3 items measure technology usage for exploitation and 3 items measure technology usage for exploration.

SCA: In a questionnaire developed by Gligor and Holcomb [14], nine items were used to measure SCA.

SSCI: Shashi et al.'s [17] questionnaire was used to measure SSCI. This questionnaire consists of 8 items; 4 items measure internal integration and 4 items measure external integration.

Firm performance: To measure firm performance, the questionnaire developed by Alipour et al. [22] was used. This questionnaire consists of 5 items.

4 RESULTS

Descriptive statistics for the variables—including mean, standard deviation, skewness, and kurtosis—are summarized in Tab. 1. A key assumption in causal modeling is multivariate normality, which requires absolute skewness values below 3 and kurtosis values under 10.

Table 1 Descriptive indices of variables

Variable	Mean	SD	Skewness	Kurtosis
Sustainable SCM	3.92	0.96	0.664	-0.617
IT	3.96	0.89	0.763	0.562
SSCI	3.81	0.85	0.594	-0.652
SCA	3.64	0.91	0.737	0.864
Firm performance	3.97	0.74	0.709	-0.671

All variables in Tab. 1 exhibit skewness and kurtosis below one. Therefore, causal modeling is normal. The matrix of correlation of variables is presented in Tab. 2.

Table 2 Correlation matrix

Variable	Sustainable SCM	IT	SSCI	SCA	Firm performance
Sustainable SCM	1				
IT	0.61**	1			
SSCI	0.51**	0.49**	1		
SCA	0.63**	0.53**	0.47**	1	
Firm performance	0.59**	0.41**	0.44**	0.56**	1

**$p < 0.01$

The data presented in Tab. 2 indicate a statistically significant correlation between the variables ($p < 0.01$). In Fig. 2, the tested model for the hypotheses is shown along with the standardized values on each path. SSCI, SCA, and firm performance are positively influenced by sustainable SCM practices.

Integrated supply chains, SCA, and firm performance are positively impacted by IT. SCA and firm performance are positively affected by SSCI. Agility in the supply chain improves the performance of firms in a significant way.

Tab. 3 reports the direct effects, t-values, and significance levels of the variables.

Table 3 Results related to direct effects

Path	β	t-value	Sig.	Variance explained
On firm performance via:				
SCA	0.30**	3.237	0.001	0.48
SSCI	0.26**	3.012	0.001	
IT	0.21**	2.784	0.001	
Sustainable SCM practices	0.44**	4.650	0.001	
On SCA via:				
SSCI	0.33**	4.638	0.001	0.55
IT	0.39**	5.041	0.001	
Sustainable SCM practices	0.49**	6.930	0.001	
On SSCI via:				
IT	0.34**	4.046	0.001	0.24
Sustainable SCM practices	0.35**	4.253	0.001	

The analysis reveals that sustainable SCM practices demonstrate significant positive associations with SSCI, SCA, and firm performance, as evidenced by the data in Tab. 3. Furthermore, IT capabilities positively influence integrated supply chain operations, SCA, and overall organizational performance. A direct positive correlation exists between SSCI and firm performance outcomes, while SCA significantly enhances firm performance metrics. The proposed model accounts for 48% of the variance in firm performance, 55% in supply chain agility, and 24% in SSCI, as detailed in Tab. 3. Another feature of SEM is to test the indirect influences of variables on each other. As shown in Tab. 4, indirect effects have been evaluated.

Table 4 Results of the indirect effect of the tested model

Path	β	t-value	Sig.
Mediating role of SSCI in the effect of:			
SCM practices on firm performance	0.09**	2.458	0.05
SCM practices on SCA	0.12**	3.135	0.01
IT on firm performance	0.09**	2.416	0.05
IT on SCA	0.11**	3.049	0.01
Mediating role of SCA in the effect of:			
SCM practices on firm performance	0.15**	2.933	0.01
IT on firm performance	0.12**	2.724	0.01
SSCI on firm performance	0.10**	2.654	0.01

The analysis in Tab. 4 reveals that SSCI serves as a significant positive mediator in four critical relationships: (1) between sustainable SCM practices and firm performance, (2) between sustainable SCM practices and SCA, (3) between IT adoption and firm performance, and (4) between IT adoption and SCA. Furthermore, the study confirms that SCA itself mediates the positive influence of IT, SSCI, and sustainable SCM on firm performance.

Goodness of fit indices of model are $GFI = 0.94$, $AGFI = 0.92$, $SRNR = 0.041$, $\chi^2/df = 1.63$, $PNFI = 0.73$, and $RMSEA = 0.054$. The results show that the obtained fit indices are all

adequate and the data has a good fit with the factor structure of this model. This methodological rigor aligns with contemporary standards for psychometric validation, as exemplified in cross-cultural adaptations of performance measurement tools by Laybidi et al. [30].

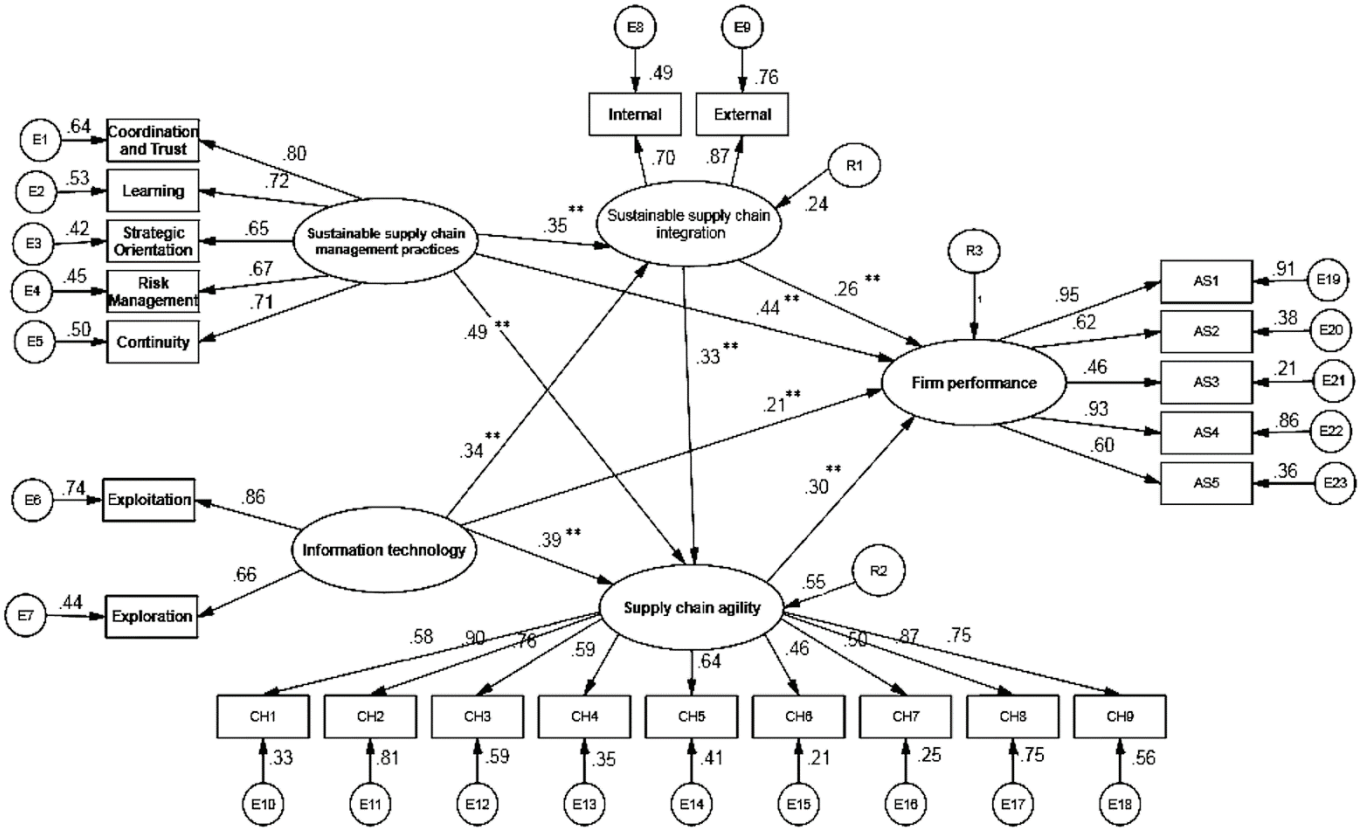


Figure 2 The tested model

5 DISCUSSION

The study considered the mediating role of agility and SSCI when modeling the influence of sustainable SCM practices and IT on firm performance. Based on the study's results demonstrated, data fit is quite good for the proposed model. It is quite capable of explaining 48% of it is variance in firm performance, 55% of it is variance in SCA, and 24% of it is variance in SSCI. In this study, the results revealed that sustainable supply chain management practices can positively affect SSCI, SCA, and firm performance in a significant manner. This finding is consistent with Hong et al. [8]. To explain this finding, SSCI and SCA will increase and ultimately firm performance will improve if the firm uses sustainable procurement methods in its supply chain process, considers its sustainability in the design of its products, uses sustainable marketing methods, evaluates the life cycle of its products, has environmental and sustainability approvals, considers purchase of the required raw materials, considers their sustainability, provides environment-related training programs for employees and suppliers, organizes environmental education activities for employees and suppliers, promotes the participation of employees and suppliers for sustainable development, and introduces environmental issues to customers and suppliers.

The impact of IT on the sustainability of supply chains, the agility of supply chains, and the performance of firms is significant and positive. This finding is consistent with Slim et al. [10]. Advanced IT infrastructures, including evolutionary data mining systems for spam detection [37], further enhance supply chain transparency by safeguarding data integrity and improving real-time decision-making efficiency. To explain this finding, SSCI and SCA will increase and ultimately firm performance will improve if the firm processes orders, issues invoices and settles with customers and provides support based on IT; information about shipping and delivery is exchanged through IT-based support with the company's primary customers, and stock and inventory information is managed through IT-based support, uses IT to understand the sales process and preferences of its primary customers, uses IT for integration of the firm and its primary customers, creates new business uses opportunities by using IT to apply firm capabilities and customer expertise, is aware of recent advancements in sciences and technologies, is able to create advanced technology processes, is able to absorb new technologies and useful innovations, has skilled tech employees in hardware and operating system, is able to master, produce or absorb key and essential commercial technology and is effective in launching programs for internal technology development or capabilities for absorbing technology from R&D centers or from suppliers and customers. The findings align with

research on CRM systems in online retailing, where IT-enabled customer engagement directly strengthens supply chain coordination and market adaptability [26]. This mirrors findings by Latifi et al. [27], where system quality in VR environments critically enhanced user engagement. Similarly, emerging IT innovations like FinTech adoption [31] and secure network infrastructures [32] demonstrate the expansive role of technology in enhancing operational efficiency across sectors. Advanced computational tools, such as the African Vulture Optimizer for reorder decision-making [28], further demonstrate how algorithmic innovations can optimize supply chain agility and operational efficiency. Emerging methodologies in big data partitioning [35] and IoT security frameworks [36] further demonstrate how advanced IT infrastructures can enhance supply chain transparency and resilience.

Integrated supply chains lead to improved SCA and performance for firms. This finding is consistent with Liu et al. [18]. According to this finding, SCA will increase and ultimately firm performance will improve if the firm is able to coordinate different teams of the organization in supply chain process, is able to coordinate tasks in different departments throughout the supply chain process, provide timely information among different departments and teams of the organization in the supply chain process, assist suppliers during the supply chain process by getting their cooperation and support, provide timely information to suppliers, is able to adapt its services to environmental changes, has adequate knowledge about characteristics of the supply chain partners, in addition to participating in SSCI, the firm's competitive position improves through interaction with supply chain partners, and supply chain partners contribute to the achievement of common goals, there is good investment in development of relationships between supply chain partners, supply chain partners use the same value system and additional efforts are made to improve relationships in the supply chain.

SCA exerts a significant positive influence on firm performance, a finding corroborated by prior research [14, 15]. This relationship can be attributed to a firm's capacity to (1) rapidly detect shifts in market dynamics, (2) identify emerging opportunities and threats, (3) streamline bidirectional information flows with suppliers and customers, (4) execute timely, data-driven decisions to mitigate risks or capitalize on opportunities, and (5) adapt operational processes—such as order specifications—to align with evolving customer demands and market conditions. Collectively, these agile capabilities enhance organizational responsiveness, driving measurable improvements in performance outcomes.

6 CONCLUSION

The findings demonstrate that SCA and SSCI act as strong mediating mechanisms, amplifying the positive effects of sustainable SCM practices on firm performance. Similarly, both agility and SSCI play critical roles in transmitting the influence of IT on organizational outcomes, with statistically significant interdependencies observed between these mediators. Notably, SCA further facilitates the relationship between SSCI and firm performance, underscoring its centrality in driving competitive advantage. These results collectively affirm that sustainable SCM practices enhance performance metrics primarily by fostering agility and integration within supply networks.

This research is constrained by its reliance on self-reported data and a sample restricted to managers, assistants, and supply chain professionals within Iranian SMEs, limiting the generalizability of conclusions. Future studies could address these gaps by employing qualitative or mixed-method approaches to explore the nuanced dynamics of sustainable SCM practices, particularly in diverse geographic and organizational contexts. Future studies could also employ nonparametric causal inference frameworks like Bayesian Additive Regression Trees (BART) to address potential nonlinear dynamics in supply chain relationships, particularly when observational data introduces unobserved heterogeneity [33, 34]. Future studies could also adopt integrated multi-criteria decision-making frameworks, such as those applied in water resource management [38], to evaluate trade-offs between environmental sustainability and operational efficiency in supply chain strategies. Additionally, triangulating self-report data with objective performance metrics (e.g., operational efficiency ratios, financial records) could strengthen the validity of findings. Future studies should adopt mixed-method approaches, combining self-reported data with spatial analyses like Social Network Analysis (SNA) to map how regional industrial clusters [23] influence supply chain resilience. Automated formative assessment tools [29] could further evaluate and enhance participant comprehension during data collection, addressing potential biases in self-reported measures. Furthermore, given evidence that monetary policy uncertainties asymmetrically influence financial market stability [39], future investigations should examine how macroeconomic policy shifts interact with supply chain agility and integration to shape firm resilience.

7 REFERENCES

[1] Rehman, S. U., Mohamed, R., & Ayoup, H. (2019). The mediating role of organizational capabilities between organizational performance and its determinants. *Journal of Global Entrepreneurship Research*, *9*(1), 1-23. https://doi.org/10.1186/s40497-019-0155-5

[2] Richard, P. J., Devinney, T. M., Yip, G. S., & Johnson, G. (2009). Measuring organizational performance: Towards methodological best practice. *Journal of Management*, *35*(3), 718-804. https://doi.org/10.1177/0149206308330560

[3] Nazari-Shirkouhi, S., Jalalat, S. M., Sangari, M. S., Sepehri, A., & Vandchali, H. R. (2023). A robust-fuzzy multi-objective optimization approach for a supplier selection and order allocation problem: Improving sustainability under uncertainty. *Computers & Industrial Engineering, 186*, 109757. https://doi.org/10.1016/j.cie.2023.109757

[4] Heidari, S., Zarei, M., Rad, S. S., Sanaei, F., Hajian, E., & Boti, M. (2025). Integrating Green Supply Chain Management and Total Quality Management: A SEM-ANN Analysis of Performance Enhancement in SMEs. *Computer and Decision Making: An International Journal, 2*, 558-569. https://doi.org/10.59543/comdem.v2i.13683

[5] Seuring, S., & Müller, M. (2008). From a literature review to a conceptual framework for sustainable supply chain

management. *Journal of cleaner production*, *16*(15), 1699-1710. https://doi.org/10.1016/j.jclepro.2008.04.020

[6] Bui, T. D., Tsai, F. M., Tseng, M. L., Tan, R. R., Yu, K. D. S., & Lim, M. K. (2021). Sustainable supply chain management towards disruption and organizational ambidexterity: A data driven analysis. *Sustainable production and consumption*, *26*, 373-410. https://doi.org/10.1016/j.spc.2020.09.017

[7] Espahbod, S., Tashakkori, A., Mohsenibeigzadeh, M., Zarei, M., Arani, G. G., Dzikuć, M., & Dzikuć, M. (2024). Blockchain-Driven Supply Chain Analytics and Sustainable Performance: Analysis Using PLS-SEM and ANFIS. *Sustainability, 16*(15), 6469. https://doi.org/10.3390/su16156469

[8] Hong, J., Zhang, Y., & Ding, M. (2018). Sustainable supply chain management practices, supply chain dynamic capabilities, and enterprise performance. *Journal of cleaner production*, *172*, 3508-3519. https://doi.org/10.1016/j.jclepro.2017.06.093

[9] Nazari-Shirkouhi, S., Miri-Nargesi, S., & Ansarinejad, A. (2017). A fuzzy decision making methodology based on fuzzy AHP and fuzzy TOPSIS with a case study for information systems outsourcing decisions. *Journal of Intelligent & Fuzzy Systems*, 32(6), 3921-3943. https://doi.org/10.3233/JIFS-12495

[10] Slim, A., Sarah, O., Kadhim, K., Ali, B., Hammood, A., & Othman, B. (2021). The effect of information technology business alignment factors on performance of SMEs. *Management Science Letters*, *11*(3), 833-842. https://doi.org/10.5267/j.msl.2020.10.019

[11] Al Humdan, E., Shi, Y., Behnia, M., & Najmaei, A. (2020). Supply chain agility: a systematic review of definitions, enablers and performance implications. *International Journal of Physical Distribution & Logistics Management*, *50*(2), 287-312. https://doi.org/10.1108/IJPDLM-06-2019-0192

[12] Nazari-Shirkouhi, S., & Samadi, S. (2025). Enhancing healthcare supply chains: A comprehensive evaluation of lean, agile, resilient and green paradigms. *Engineering Applications of Artificial Intelligence, 145*, 110204. https://doi.org/10.1016/j.engappai.2025.110204

[13] Hines, P., Holweg, M., & Rich, N. (2004). Learning to evolve: A review of contemporary lean thinking. *International Journal of Operations & Production Management, 24*(10), 994-1011. https://doi.org/10.1108/01443570410558049

[14] Gligor, D. M., & Holcomb, M. C. (2012). Antecedents and consequences of supply chain agility: establishing the link to firm performance. *Journal of Business logistics*, *33*(4), 295-308. https://doi.org/10.1111/jbl.12003

[15] Zhu, M., & Gao, H. (2021). The antecedents of supply chain agility and their effect on business performance: an organizational strategy perspective. *Operations Management Research*, *14*(1), 166-176. https://doi.org/10.1007/s12063-020-00174-9

[16] Nazari-Shirkouhi, S., Keramati, A., & Rezaie, K. (2015). Investigating the effects of customer relationship management and supplier relationship management on new product development. *Tehnički vjesnik*, 22(1), 191-200. https://doi.org/10.17559/TV-20140623130536

[17] Shashi, S., Cerchione, R., Centobelli, P., & Shabani, A. (2018). Sustainability orientation, supply chain integration, and SMEs performance: A causal analysis. *Benchmarking: An International Journal*, *25*(9), 3679-3701. https://doi.org/10.1108/BIJ-08-2017-0236

[18] Liu, A., Liu, H., & Gu, J. (2021). Linking business model design and operational performance: The mediating role of supply chain integration. *Industrial Marketing Management*, *96*, 60-70. https://doi.org/10.1016/j.indmarman.2021.04.009

[19] Jiang, Q., & McCabe, S. (2021). Information technology and destination performance: Examining the role of dynamic capabilities. *Annals of Tourism Research*, *91*, 103292. https://doi.org/10.1016/j.annals.2021.103292

[20] Nazari-Shirkouhi, S., Badizadeh, A., Dashtpeyma, M., & Ghodsi, R. (2023). A model to improve user acceptance of e-services in healthcare systems based on technology acceptance model: an empirical study. *Journal of Ambient Intelligence and Humanized Computing, 14*(6), 7919-7935. https://doi.org/10.1007/s12652-023-04601-0

[21] Nasiri, S., Shahabi, S., Shafiesabet, A., Talebbeidokhti, M., & Behineh, E. A. (2026). Cybersecurity in Action: Unraveling the Effects of Individual, Social, and Organizational Determinants. *Tehnički glasnik, 20*(2). Accepted but not published yet. https://doi.org/10.31803/tg-20240627004731

[22] Alipour, N., Nazari-Shirkouhi, S., Sangari, M. S., & Vandchali, H. R. (2022). Lean, Agile, Resilient, and Green Human Resource Management: The Impact on Organizational Innovation and Organizational Performance. *Environ Sci Pollut Res, 29*, 82812–82826. https://doi.org/10.1007/s11356-022-21576-1

[23] Bevilacqua, C., Sohrabi, P., & Hamdy, N. (2022). Spatializing Social Networking Analysis to Capture Local Innovation Flows towards Inclusive Transition. *Sustainability, 14*(5), 3000. https://doi.org/10.3390/su14053000

[24] Saremi, Y., Aghaei, M., Esfid, M. R., Hasangholipour, T., & Ansari, M. (2024). Providing Digital Marketing Model with Customer Participation Approach with intervening role of knowledge management. *International Journal of Knowledge Processing Studies, 4*(2), 148-167. https://doi.org/10.22034/kps.2023.410230.1154

[25] Azimi Asmaroud, S. (2022). Preservice Elementary Teachers' Categorical Reasoning and Knowledge Transfer on Definition Tasks with Two Dimensional Figures. *Theses and Dissertations*. 1588. https://ir.library.illinoisstate.edu/etd/1588

[26] Latifi, K., Ebrahimi, A., Ranjbaran, M., Mirzaei, A., & Fakhri, Z. (2023). Efficient customer relationship management systems for online retailing: The investigation of the influential factors. *Journal of Management & Organization, 29*(4), 763-798. https://doi.org/10.1017/jmo.2022.65

[27] Latifi, T., Li, J., Blum, S. C., & Fowler, D. (2024). Determinants of Users' Intention to Visit a Destination: A Virtual Reality Quality Framework. *Journal of Quality Assurance in Hospitality & Tourism*, 1-25. https://doi.org/10.1080/1528008X.2024.2440010

[28] Bahadoran Baghbadorani, S., Johari, S. A., Fakhri, Z., Khaksar Shahmirzadi, E., Navruzbek Shavkatovich, S., & Lee, S. (2022). A new version of african vulture optimizer for apparel supply chain management based on reorder decision-making. *Sustainability, 15*(1), 400. https://doi.org/10.3390/su15010400

[29] Karizaki, M. S., Gnesdilow, D., Puntambekar, S., & Passonneau, R. J. (2024, July). How Well Can You Articulate that Idea? Insights from Automated Formative Assessment. In *International Conference on Artificial Intelligence in Education* (pp. 225-233). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-64299-9_16

[30] Laybidi, M. I., Asghari-Jafarabadi, M., Maleki-Ghahfarokhi, A., Entezarizarch, E., & Danesh, M. K. (2025). Translation and Cross-Cultural Adaptation of the Individual Work Performance Questionnaire (IWPQ) into Persian and Assessing its Psychometric Properties. *Journal of Health and Safety at Work, 15*(1), 125-143.

[31] Rivandi, E. (2024). FinTech and the Level of Its Adoption in Different Countries around the World. Available at *SSRN* 5049827. https://doi.org/10.2139/ssrn.5049827

[32] Rivandi, E., & Jamili Oskouie, R. (2024). A Novel Approach for Developing Intrusion Detection Systems in Mobile Social Networks. Available at *SSRN* 5174811. https://doi.org/10.2139/ssrn.5174811

[33] Mahdavi, P., Ehsani, M. A., Ahelegbey, D., & Mohammadpour, M. (2023). Measuring Causal Effect with ARDL-BART: A Macroeconomic Application. *Applied Mathematics, 15*(4), 292-312. https://doi.org/10.4236/am.2024.154018

[34] Mahdavi, P., & Ehsani, M. A. (2023). Dynamic Causal Effects in Econometrics with a Focus on the Nonparametric Method: A Review Paper. *Iranian Journal of Economic Studies, 11*(2), 427-449. https://doi.org/10.22099/ijes.2023.44459.1885

[35] Soltanmohammadi, E., Dilek, A., & Hikmet, N. (2024). Tailored Partitioning for Healthcare Big Data: A Novel Technique for Efficient Data Management and Hash Retrieval in RDBMS Relational Architectures. *Journal of Data Analysis and Information Processing, 13*(1), 46-65. https://doi.org/10.4236/jdaip.2025.131003

[36] Khatami, S. S., Shoeibi, M., Oskouei, A. E., Martín, D., & Dashliboroun, M. K. (2025). 5DGWO-GAN: A Novel Five-Dimensional Gray Wolf Optimizer for Generative Adversarial Network-Enabled Intrusion Detection in IoT Systems. *Computers, Materials and Continua, 82*(1), 881-911. https://doi.org/10.32604/cmc.2024.059999

[37] Ehsani Chimeh, H., & Karami, M. (2018). Spam detection from big data based on evolutionary data mining systems. *Transactions on Machine Intelligence*, 1(1), 1-9.

[38] Fathi, S., Bahrami, J., Momeni, S., Seyedrezaei, M., & Javidi Sabbaghian, R. (2025). A Multi-Criteria Framework for Water Resource Management: Combining WEAP and TOPSIS for Optimal Solutions. *Advances in Civil Engineering and Environmental Science, 2*(1), 79-91. https://doi.org/10.22034/acees.2025.484672.1011

[39] Roshdieh, N. (2024). The Effect of Monetary Policy Uncertainty on Stock Market Uncertainty with NARDL Approach. *Research Journal of Finance and Accounting, 15*(10), 1-9. https://doi.org/10.2139/ssrn.5110149

Authors' contacts:

Asiye Omidian
Department of Economics, University of Missouri-Columbia,
615 Locust St, Columbia, MO 65201, USA

Athena Omidian
(Corresponding author)
Department of Marketing, Carl H. Lindner College of Business,
University of Cincinnati,
Cincinnati, OH, USA
omidiaaa@mail.uc.edu

Shahriar Mohajer
Department of Research & Development, Supply Chain Unit, Kermanshah Branch,
Alborz Paradise Export Development Company (APED.global), Kermanshah, Iran

Elnaz Zarei
Department of Information Technology Management—Comprehensive Quality System, Yazd University of Science and Art, Iran

Azita Ramezani
Applied Data Science, Cleveland State Cleveland State University Cleveland,
2121 Euclid Ave, Cleveland, OH 44115, USA

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250325111250

Original scientific paper
Received: 2025-03-25, Accepted: 2025-09-02

Hierarchical Ensemble Learning for Automatic Classification of Construction Work Codes

Yeong-Chae Yun, Seok-Heon Yun*

Abstract: The Work Classification Code (WCC) serves as a critical tool for standardizing Bill of Quantities (BOQ) in construction projects, ensuring reliable cost estimation and effective expense management. However, current BOQ–WCC matching processes rely heavily on subjective judgment, resulting in inconsistencies and limited standardization. To address this issue, this study constructed a dataset of 8,021 valid items collected from six school construction BOQs commissioned by the Korea Public Procurement Service (PPS), and developed an automated matching model using hierarchical ensemble learning. The proposed approach reflects the five-level hierarchical structure of the WCC (Levels 1–5) and applies different ensemble methods according to classification complexity: Bagging at Levels 1–2, Boosting at Levels 3–4, and Weighted Voting at Level 5. A baseline comparison with a single LSTM model was also conducted, confirming that ensemble methods consistently outperformed the base model, particularly at intermediate levels where classification complexity is higher. Experimental results demonstrated that Bagging provided stable improvements at upper levels, Boosting achieved substantial gains at intermediate levels, while Weighted Voting offered limited benefits at the lowest level. Despite these improvements, overall performance declined as class granularity increased, and Level 5 classification showed severe degradation due to extreme data sparsity. This study confirms the feasibility of applying hierarchical ensemble learning for automated WCC matching and highlights its potential for improving BOQ standardization in real-world projects. The findings suggest that automatic WCC assignment can support reliable cost reviews and enable retroactive coding of BOQs where codes were previously absent, thereby enhancing efficiency in design-phase verification systems. Future work will focus on addressing data sparsity through large-scale BOQ collection, incorporating advanced pre-trained language models such as BERT, refining the hierarchical code structure through data integration, and validating model applicability in public procurement and contractor systems.

Keywords: Automatic Code Matching; Bill of Quantities; Construction Work Classification Code; Ensemble Learning; Hierarchical Classification; Machine Learning

1 INTRODUCTION

In construction projects, the Bill of Quantities (BOQ) is the key document for cost estimation and contract administration. The BOQ systematically lists item descriptions, specifications, units, quantities, unit prices, and the total construction cost, and it is prepared on the basis of design drawings and specifications; thus, it serves as the benchmark for deriving construction costs [1, 2]. This enables cost review and comparison, and—particularly in public projects—clarifies contract terms and enhances the reliability of cost management. To distinguish trades and items, BOQs are accompanied by construction codes, which play an important role in item identification and cost organization. However, the construction codes currently in use vary across agencies and projects, standardization is insufficient, and even identical trades may be expressed differently; moreover, subjective judgment by the preparer is easily introduced [1, 3]. As a result, inter-project cost comparisons become difficult and data consistency cannot be ensured.

To address these issues, Korea's Public Procurement Service provides the Work Classification Code (WCC), which guides the preparation of standardized BOQs. Through a 12-digit hierarchical structure, the WCC systematically classifies trades, standardizes BOQ composition, and facilitates cost comparison and review across projects [4]. Nevertheless, limited practitioner understanding of the WCC's structure and application often leads to the continued use of legacy practices, constraining its uptake in practice.

The purpose of this study is to develop an algorithm for automatically matching BOQ items with the standard WCC. To this end, a hierarchical classification model based on ensemble learning was designed, incorporating Bagging (Bootstrap Aggregating), Boosting, and Weighted Voting techniques at different levels to evaluate performance. The aim is to support BOQ standardization, enhance the reliability of cost review, and reduce the complexity of practitioners' work processes.

The remainder of this paper is structured as follows. First, prior research related to the WCC system is reviewed, and the structure and hierarchical characteristics of the WCC are described. The methodology and hierarchical ensemble model are then presented, followed by an analysis of the experimental results. Finally, the study's contributions and directions for future research are discussed.

2 LITERATURE REVIEW

To systematically review construction costs and manage project expenses, the Korea Public Procurement Service introduced the Construction Work Classification Code (WCC). WCC adopts a hierarchical classification structure, enabling precise categorization of construction work items and ensuring consistency in bill of quantities (BOQ) preparation [4]. According to Lee et al. [4], the application of WCC enhances cost review, supports comparative analysis, and provides a more systematic framework for managing cost variations across trades. However, challenges remain in effectively adopting WCC in practice [5].

One key reason for this limited adoption is the complexity of the coding structure. The WCC consists of a 12-digit hierarchical system, with distinct classification rules at each level. Practitioners often lack sufficient understanding of these rules and continue to prepare BOQs in traditional ways, leaving room for subjective judgment in code selection. Moreover, identical work items may be represented differently across projects, leading to inconsistencies in code application. Yun [6] also emphasized that the absence of a fully standardized coding system has resulted in institutions and designers relying on different construction codes, undermining interoperability.

This standardization challenge is not unique to Korea. In North America, databases based on MasterFormat and

OmniClass have been established, with Guven et al. [7] highlighting the necessity of systematic classification to track and manage construction resources. In the UK, Pupeikis et al. [8] (2022) compared Uniclass 2015 and CCI, noting both the potential of hierarchical classification systems in BIM environments and the practical challenges of implementation. Similarly, Danusevics et al. [9] (2022) conducted a comparative analysis of more than 20 national classification systems, emphasizing that while standardized coding offers clear advantages, its adoption and challenges vary across countries, underscoring the importance of international harmonization. In Spain, policy-level initiatives have begun requiring the integration of BIM-based classification systems and a Common Data Environment (CDE) in public procurement, signaling institutional efforts to enhance consistency in BOQ preparation and data management [10]. Collectively, these international cases demonstrate that the classification and standardization of construction work items are universal challenges across the construction industry.

At the same time, the BOQ consists of textual data representing diverse work items, making effective text classification essential for structuring specifications, organizing documents, and supporting cost estimation [11]. Given the variety of expressions that may be used in BOQ descriptions, recent studies have increasingly applied natural language processing (NLP) techniques to analyze construction data. For example, Shamshiri, M., Chi, S., & Moon, S. [12] reported the application of NLP to construction document analysis, while Zhang et al. [13] explored the use of text mining for automated classification of construction-related data. These studies demonstrate the potential for automation but often fail to fully incorporate hierarchical code structures or remain limited to specific datasets.

Against this backdrop, this study positions the BOQ–WCC matching task as a hierarchical text classification problem and applies an ensemble-based approach to systematically evaluate performance improvements.

3 STANDARD CONSTRUCTION WORK CLASSIFICATION CODE STRUCTURE

In Korea, construction work codes were developed to standardize BOQ by benchmarking UniClass from Europe and MasterFormat from North America. The standard WCC is a construction information classification system established by the Korea Public Procurement Service to systematically organize construction items in BOQ and standardize cost estimation. The WCC consists of a 12-digit alphanumeric code, where each digit carries a specific meaning, reflecting the hierarchical structure of the classification [6].

The first digit of the WCC represents the type of construction work. The second to sixth digits classify the work into major, medium, minor, detailed, and sub-detailed categories, distinguishing major work types, item names, and specifications. The seventh and subsequent digits provide additional information to specify particular materials or configurations. This hierarchical structure allows the WCC to further refine construction work classification from broader categories at the top level to more detailed categories at the lower levels.

As shown in the Tab. 1, the number of classes increases exponentially as the level deepens. In particular, Levels 4 and 5 primarily capture specification-level details, where data becomes sparse and highly imbalanced.

Table 1 Summarizes the number of classes at each level in the full WCC system

Level	Definition	Number of Classes
Level 1	Major work type	20
Level 2	Intermediate work type	131
Level 3	Minor work type	389
Level 4	Detailed work type	939
Level 5	Sub-detailed classification (e.g., specifications)	1,603

As shown in Fig. 1, for the purposes of this study, only the second to sixth digits of the WCC were used, corresponding to Levels 1 through 5. The first digit, which simply distinguishes between broad categories of construction work, was excluded. This approach ensures that the hierarchical structure of the WCC is maintained while optimizing analytical efficiency, as most construction items can be effectively categorized within this five-digit framework. Importantly, item identification in actual BOQs is most commonly performed up to Level 3, whereas Levels 4 and 5 represent detailed specification-level distinctions. Accordingly, performance analysis in this study was conducted with these structural characteristics in mind.

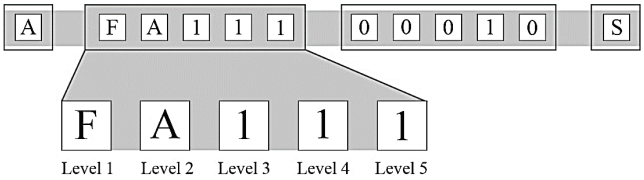


Figure 1 Structure of the Standard Construction Work Classification Code and Selected 5-Digit Levels for Analysis

4 METHODOLOGY

4.1 Dataset and Preprocessing

For this study, six BOQs from school construction projects commissioned by the Korea Public Procurement Service (PPS) were collected to construct the dataset for automatic matching between the Standard Work Classification Code (WCC) and BOQ items. The original BOQs contained approximately 12,000 entries, and after removing duplicates, 8,021 valid items were retained. Each item consists of textual information such as the major category, item description, and specifications, along with the corresponding WCC label. As shown in Tab. 2, for instance, an item categorized under Plastering Works with the description Mortar Plastering is associated with the code GA112.

When classified by major category, the dataset exhibited noticeable imbalance. Categories such as Windows and Glass Works and Metal Works accounted for more than 20% of the dataset, whereas others such as Aggregate Costs, Transportation Costs, and Demolition Works represented less than 1%. Such imbalance increases the risk of overfitting toward majority classes and deteriorates classification performance for minority categories.

Table 2 Example of Hierarchical Classification of Standard WCC

Major Category	Item Description	Level 1	Level 2	Level 3	Level 4	Level 5
Plastering Works	Mortar Plastering	G	GA	GA1	GA11	GA112
Waterproofing Works	Cement Liquid Waterproofing	H	HI	HI0	HI00	HI000
Masonry Works	0.5B Bricklaying	F	FA	FA1	FA11	FA111

Table 3 Distribution of Valid Dataset by Standard Work Categories

Work Category	Count	Percentage (%)
Window and Glass Works	2,079	25.79
Metal Works	810	10.05
Common Works	569	7.06
Finishing Works	522	6.47
Plastering Works	498	6.18
Waterproofing Works	487	6.04
Temporary Works	401	4.97
Reinforced Concrete Works	376	4.66
Masonry Works	354	4.39
Ancillary Works	338	4.19
Carpentry Works	334	4.14
Tile and Stone Works	314	3.89
Painting Works	305	3.78
Roof and Gutter Works	176	2.18
Foundation and Substructure Works	148	1.84
Miscellaneous Works	136	1.69
Earth Works	114	1.41
Demolition Works	66	0.82
Transportation Cost	20	0.25
Aggregate Cost	15	0.19
Total	**8,021**	**100**

To address this issue, the Synthetic Minority Over-sampling Technique (SMOTE) was applied. Unlike simple duplication, SMOTE generates synthetic samples by interpolating between existing minority class samples, thereby balancing the data distribution more effectively. Before applying SMOTE, some categories accounted for less than 5% of the dataset; after application, the minimum share increased to approximately 8–10%. This adjustment mitigated class imbalance and reduced training bias.

4.2 Research Workflow and Hierarchical Ensemble Model Design

The proposed methodology followed four sequential stages: data collection and preprocessing, feature extraction, hierarchical ensemble model training, and performance evaluation. The entire workflow, from dataset construction to results analysis, is illustrated in Fig. 2.

The Standard WCC is hierarchically structured, with categories becoming increasingly granular from higher to lower levels. Thus, rather than employing a single flat classifier, a hierarchical approach was adopted to reflect varying levels of complexity. Accordingly, the classification was performed progressively from Level 1 to Level 5.

Text preprocessing is a critical step in determining model quality [14]. In this study, the input features were constructed by combining major category and item description. A synonym dictionary was applied to normalize equivalent terms, and unnecessary symbols and whitespace were removed. The Keras Tokenizer was used to limit the vocabulary size to 5,000 tokens, after which texts were converted into integer sequences and padded to a maximum length of 26 tokens. For text vectorization, a Word2Vec-based embedding layer was applied, generating 128-dimensional embedding vectors. The embedding weights were updated during training, allowing the model to learn distributed representations optimized for BOQ text. Compared to frequency-based representations, this embedding approach more effectively captured semantic similarities between terms. The resulting embedding vectors were then fed into an LSTM network to capture sequential context, after which ensemble methods were applied according to the characteristics of each classification level.

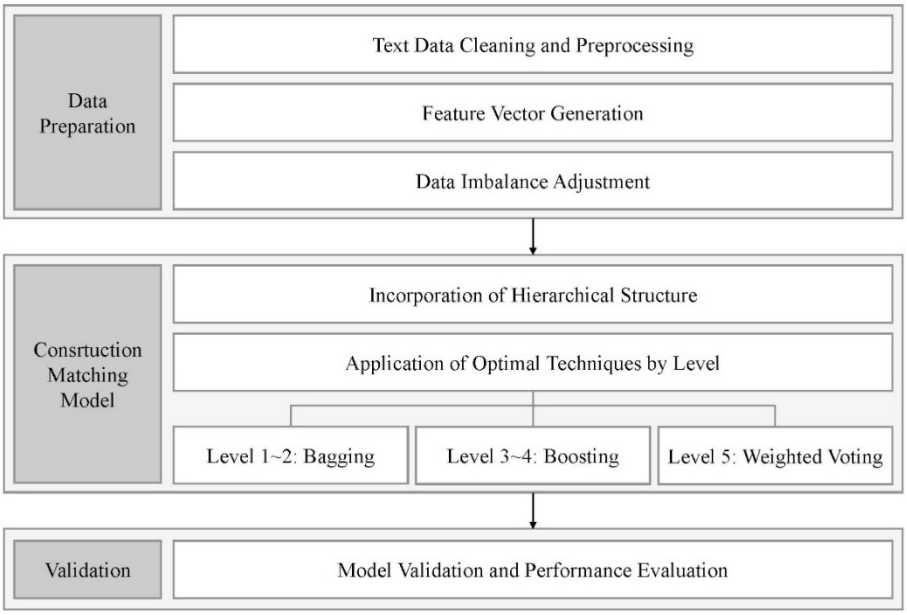


Figure 2 Framework of the Hierarchical Construction Code Matching Model

4.3 Bagging-Based Model

Bagging (Bootstrap Aggregating) is an ensemble learning technique that trains multiple predictive models and derives the final prediction through majority voting (for classification tasks) or averaging (for regression tasks). According to Breiman et al. [15], bagging reduces variance and enhances model stability by ensuring that individual models are trained independently.

In this study, bagging was applied to Level 1 and Level 2. At these upper levels, the classification scope is broader, and the dataset contains a larger volume of samples, allowing individual models to learn diverse patterns and achieve stable performance. To implement this, a Random Forest-based bagging model was employed. Random Forest operates by training multiple decision trees and aggregating their predictions to generate the final output [16].

The Random Forest-based bagging approach ensures robust predictions by mitigating data variability and outperforming single models in terms of generalization. Additionally, since each model within the ensemble is trained independently, bagging effectively prevents overfitting, further contributing to classification accuracy at Level 1 and Level 2.

4.4 Boosting-Based Model

Boosting is an ensemble learning technique that sequentially trains multiple weak learners to progressively enhance predictive performance. According to Yaman et al. [17], boosting corrects the errors of previous models, making it an effective approach for constructing a highly accurate predictive model.

In this study, boosting was applied to Level 3 and Level 4. At these lower levels, construction work classification becomes more granular, increasing the complexity of classification. Therefore, boosting was deemed appropriate as it incrementally improves the learning performance of individual models. To implement this, the XGBoost (eXtreme Gradient Boosting) algorithm was employed to develop the boosting-based model. XGBoost follows the gradient boosting framework, where misclassified samples from previous iterations are assigned higher weights so that subsequent models focus on correcting these errors [18]. In this study, the boosting-based model achieved a higher F1-score compared to individual models, demonstrating improved classification performance.

Since boosting adapts to data characteristics during training, it is particularly effective for handling higher classification complexity at Level 3 and Level 4. However, boosting models are prone to overfitting, so regularization parameters were adjusted to optimize model complexity and prevent excessive variance.

4.5 Weighted Voting-Based Model

Voting is an ensemble learning technique that combines multiple models to generate the final prediction. In majority voting (MV), all classifiers are assigned equal weights, whereas weighted voting (WV) adjusts model-specific weights based on their reliability [19]. According to Livieris et al. [20], weighted voting is an effective technique that leverages the predictive power of individual classifiers to enhance classification accuracy.

In this study, weighted voting was applied at Level 5, the most granular classification stage in the WCC hierarchy. At this level, the reliability of individual model predictions tends to decline due to the complexity and sparsity of the data. Therefore, an ensemble of multiple models was trained, and the final prediction was made using weighted voting, with the assigned weights reflecting each model's reliability. To assign weights, an F1-score-based weighting strategy was adopted. Specifically, models with higher F1-scores were assigned greater weights according to performance evaluation metrics. In this study, bagging, boosting, and logistic regression models were combined, and their weights were adjusted proportionally to their F1-scores to optimize prediction performance. Experimental results showed that the weighted voting approach achieved a higher F1-score than individual models, effectively compensating for performance discrepancies among different classifiers. Weighted voting is particularly effective in scenarios where data sparsity is high and performance variance among individual models is significant. By implementing weighted voting at Level 5, this study aimed to enhance model reliability and achieve more accurate construction code matching results.

5 RESULTS AND DISCUSSION

In this study, a hierarchical ensemble model was applied to perform automatic matching between the standard WCC and BOQ items. The model's performance was evaluated at each level using F1-score, Precision, and Recall. Additionally, the early stopping points were analyzed to determine the optimal epoch values at which the model weights were saved.

5.1 Performance Evaluation of Level 1 and Level 2

At Levels 1 and 2, a bagging-based ensemble method was applied. Tab. 4 presents the performance evaluation results, where F1-score, Precision, and Recall were used as the primary assessment metrics. At Level 1, the model achieved an F1-score of 0.9805, Precision of 0.9869, and Recall of 0.9762, demonstrating highly accurate classification performance. Early stopping occurred at epoch 16, with the optimal model weights saved at epoch 13, indicating that the model completed training at an appropriate point. The high performance at Level 1 can be attributed to clear distinctions between classes and the absence of severe class imbalance issues. In addition, the relatively balanced class distribution at this level allowed the bagging method to stabilize learning by reducing variance across individual learners, which contributed to consistent improvements in precision and recall.

Similarly, at Level 2, the bagging-based ensemble learning approach was employed, achieving an F1-score of 0.8728, Precision of 0.8799, and Recall of 0.8747. Early

stopping was triggered at epochs 8 and 11, with the optimal weights stored at epochs 5 and 8. While the performance was slightly lower than at Level 1, this decline is likely due to the increased number of classes that needed to be classified. Despite the decrease, the application of bagging reduced the risk of overfitting, which is often observed in single models at this stage, thereby maintaining stable performance even with moderately increased complexity.

Table 4 Level 1 and Level 2 Performance Metrics

Level	F1-score	Precision	Recall
1	0.981	0.987	0.976
2	0.873	0.880	0.875

5.2 Performance Evaluation of Level 3 and Level 4

At Levels 3 and 4, a boosting-based ensemble method was applied to evaluate classification performance. As shown in Tab. 5, Level 3 achieved an F1-score of 0.7423, Precision of 0.7496, and Recall of 0.7549. At this level, the number of classes increased to 548, significantly increasing classification complexity compared to Levels 1 and 2. However, the model maintained relatively stable performance, which can be attributed to the error correction mechanism of the boosting technique. Since boosting improves predictions by iteratively correcting errors from previous models, it effectively retained high classification accuracy even under complex classification conditions. Early stopping occurred at epoch 14, with the optimal weights saved at epoch 11, indicating that the model converged at an appropriate point. Nevertheless, the class distribution at Level 3 showed moderate imbalance, and boosting's ability to focus on misclassified samples proved advantageous in mitigating the underperformance of minority classes.

Similarly, at Level 4, the boosting-based ensemble method was employed, achieving an F1-score of 0.4655, Precision of 0.4645, and Recall of 0.4991. At this level, the number of classes increased to 1,490, approximately three times that of Level 3, making classification significantly more challenging. Performance declined considerably due to the increased complexity, and the model struggled to effectively learn certain classes. Early stopping occurred at epoch 9, with the optimal weights saved at epoch 6, suggesting that while the model reached optimal performance in a relatively short training period, performance degradation was observed as classification difficulty increased. The sharp drop in performance can be attributed not only to the explosion in the number of classes but also to severe class imbalance, where many categories were underrepresented. This indicates that boosting alone was insufficient to fully overcome the sparsity problem at deeper levels.

Table 5 Level 3 and Level 4 Performance Metrics

Level	F1-score	Precision	Recall
3	0.742	0.750	0.755
4	0.466	0.465	0.499

5.3 Performance Evaluation of Level 5

At Level 5, the weighted voting-based ensemble method was applied. As shown in Tab. 6, the model achieved an F1-score of 0.0001, Precision of 0.0001, and Recall of 0.0011. Early stopping occurred at epoch 4, with the optimal weights stored at epoch 1. The model encountered increasing classification complexity due to the large number of classes and insufficient samples per class, making it difficult to distinguish between classes effectively. Consequently, learning was not successfully completed, and the model struggled to capture meaningful patterns.

Furthermore, despite applying the weighted voting approach, the low performance of individual models limited its effectiveness. Weighted voting aggregates predictions by assigning higher weights to more reliable models, but at Level 5, individual models failed to achieve meaningful accuracy, making it difficult to improve final predictions through weight adjustments.

During training, validation accuracy remained close to zero, indicating a severe imbalance between training and validation data. Additionally, early stopping occurred at an initial stage, preventing sufficient learning, and the model exhibited rapid convergence without significant loss reduction during training.

This near-zero performance at Level 5 should be regarded as a critical limitation of the proposed model, reflecting the extreme sparsity and over-fragmentation of classes at this level. However, since most BOQ classifications in practice are determined within Levels 1–3, the limited applicability of Level 5 does not invalidate the practical contribution of this study. Instead, it highlights the need for future research to explore larger datasets, external specification libraries, or transfer learning techniques to address this bottleneck.

Table 6 Level 5 Performance Metrics

Level	F1-score	Precision	Recall
5	0.0001	0.0001	0.0011

5.4 Summary Analysis

The overall comparison of F1-score, Precision, and Recall across all levels is shown in Fig. 3. Overall, all performance metrics exhibit a gradual decline as the level increases. At Level 1, the model demonstrates excellent performance with an F1-score of 0.981, Precision of 0.987, and Recall of 0.976, indicating stable and reliable classification. At Level 2, the performance remains strong with an F1-score of 0.873, Precision of 0.880, and Recall of 0.875. However, a noticeable decline occurs at Level 3, where the F1-score drops to 0.742, Precision to 0.750, and Recall to 0.755. A further significant degradation is observed at Level 4, with an F1-score of 0.466, and at Level 5 the performance nearly collapses, showing that the model struggles to effectively distinguish fine-grained categories.

In addition, to validate the effectiveness of the proposed ensemble approach, a baseline comparison with a single LSTM model was conducted using the same dataset of 8,021

items. The overall comparison of F1-score changes of ensemble models relative to the base LSTM model across all levels is presented in Tab. 7. The results confirmed that ensemble methods consistently outperformed the single model across most levels, although the degree of improvement varied depending on classification complexity. Bagging provided moderate but stable improvements, with gains of approximately 3–5% at Levels 1–2 and up to 10% at Levels 3–4, but only marginal effects at Level 5. Boosting achieved the most substantial improvements, with F1-score increases of about 15% at Level 1, 18% at Level 2, and up to 20% at Level 3, although its effectiveness diminished at Level 5 due to severe class sparsity. Weighted Voting, by contrast, did not yield consistent gains and in some cases underperformed relative to the baseline, particularly at Level 5 where extreme data sparsity remained a critical challenge.

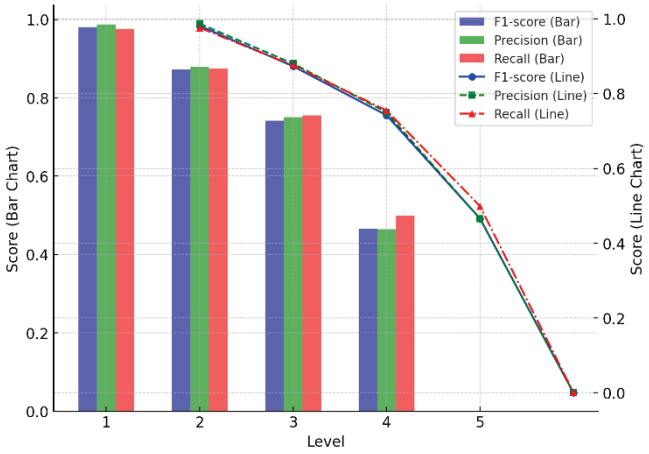


Figure 3 Comparison of F1-score, Precision, and Recall by Level

Table 7 F1-score changes of ensemble models relative to the base LSTM model

Ensemble method	Level 1 (%)	Level 2 (%)	Level 3 (%)	Level 4 (%)	Level 5 (%)
Bagging	3.5	4.2	2.0	1.5	0.0
Boosting	5.0	6.0	15.0	12.0	0.0
Weighted Voting	0.5	0.3	–0.5	–0.8	0.0

These findings demonstrate that while hierarchical ensemble learning is effective in handling multi-level construction code classification, its benefits are uneven across levels. The upper and middle levels (Levels 1–4) showed notable improvements compared to the single model, whereas Level 5 performance remained limited, highlighting the need for additional methodological enhancements and more robust datasets to overcome data sparsity in fine-grained classification.

From a practical perspective, these findings suggest that the hierarchical ensemble model can serve as a useful decision-support tool in construction project management. In particular, the stable performance observed at Levels 1–3 indicates that the model can reliably support the classification of major and medium-level work categories, which are the levels most frequently referenced in cost estimation and project planning. However, the near-zero performance at Level 5 highlights a significant limitation: when applied directly to detailed specifications or highly granular categories, the current approach may fail to provide meaningful guidance. This implies that while the proposed method has potential for practical adoption in tasks such as BOQ standardization and cost control, additional methodological enhancements and larger, more balanced datasets are required to ensure robustness at the most detailed levels.

6 CONCLUSION

To address this issue, this study aimed to develop an automated matching model between BOQ items and the standard WCC using advanced machine learning techniques. Additionally, a hierarchical ensemble learning approach was proposed to incorporate the structured characteristics of the standard WCC into the model.

In this study, a hierarchical classification model was constructed using Bagging, Boosting, and Weighted Voting techniques. Bagging-based Random Forest models were applied at Levels 1 and 2, ensuring stable classification performance. Boosting-based XGBoost models were employed at Levels 3 and 4 to improve classification accuracy in more complex classification tasks. At Level 5, Weighted Voting was implemented to aggregate model predictions based on individual model reliability.

Each ensemble method demonstrated varying performance depending on data complexity and classification level. The experimental results showed high classification accuracy at Level 1 (F1-score: 0.9805) and Level 2 (F1-score: 0.8728). Level 3 maintained relatively stable performance (F1-score: 0.7423), but performance degradation was observed at Level 4 (F1-score: 0.4655) due to increased classification complexity. Level 5 exhibited a significant decline in accuracy (F1-score: 0.0001), indicating that classification performance deteriorated as the number of classes increased and data sparsity intensified.

These findings highlight the importance of selecting appropriate ensemble techniques at different levels to optimize model performance while also emphasizing the need for improvement at lower levels. The hierarchical ensemble learning approach used in this study demonstrated high practical applicability at higher levels, particularly with Bagging, which provided stable predictions. However, as the number of classes increased at Level 3 and beyond, performance declined due to overfitting and data imbalance. Level 5 faced severe data sparsity issues, limiting the effectiveness of Weighted Voting in improving classification accuracy.

From a practical perspective, however, Levels 1–3 cover the majority of BOQ classification tasks in real projects, meaning that the proposed model can already be used to automatically assign codes to uncoded BOQ items and assist in cost verification during the design stage. Nonetheless, the near-zero performance at Level 5 represents a critical research limitation. Although its practical impact is relatively minor, as fine-grained specifications are less frequently required in BOQ preparation, overcoming this limitation remains essential for the long-term advancement of fully automated classification.

Future work will therefore prioritize: (1) expanding the dataset with larger and more diverse BOQ samples, (2) applying pre-trained embeddings such as BERT or extended Word2Vec for richer semantic representation, (3) refining the WCC hierarchy and integrating external datasets to enhance coverage, and (4) validating the model's deployment within real systems used by public agencies and contractors. By addressing these aspects, the proposed model can evolve into a fully operational tool for automated BOQ standardization and cost management, ultimately enhancing transparency and efficiency in construction project delivery.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024-00338908).

7 REFERENCES

[1] Noh, H. R., & Yun, S. H. (2017). A case study on the educational facility project for the improvement of work item's structure in BoQ. *Journal of the Architectural Institute of Korea Structure & Construction*, *33*(8), 47–54. https://doi.org/10.5659/JAIK_SC.2017.33.8.47

[2] Bikçe, M., & Göçer, S. (2018). An approach to preparing a bill of quantities. *International Advanced Research Engineering Journal*, *2*(3), 229–233. Available from https://dergipark.org.tr/en/pub/iarej/issue/40961/407726

[3] Cho, H. H., Kang, T. K., Lee, Y. S., & Cho, M. Y. (1999). A study on the reformation of the BoQ structure for public building projects in Korea. *Journal of the Architectural Institute of Korea Structure & Construction*, *15*(9), 123–131. Available from https://www.riss.kr/link?id=A3397343

[4] Lee, G. E., Lee, G. T., Chi, S. H., & Oh, S. O. (2023). Automatic classification of construction work codes in bill of quantities of national roadway based on text analysis. *Journal of Construction Engineering and Management*, *149*(2), 04022163. https://doi.org/10.1061/JCEMD4.COENG-12730

[5] Park, H. P., & Lee, J. S. (2004). A promotion plan through measuring the utilization of information classification systems in the construction industry. *Korean Journal of Construction Engineering and Management*, *5*(6), 90–100. UCI: G704-001084.2004.5.6.021

[6] Yun, S. H. (2010). A study on the standardized cost code system for BoQ for efficiency of cost management in public construction projects. *Journal of the Architectural Institute of Korea Structure & Construction*, *26*(12), 167–174. Available from https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART001506512

[7] Guven, G., Arceo, A., Bennett, A., Tham, M., Olanrewaju, B., McGrail, M., Isin, K., & Olson, A. (2022). A construction classification system database for understanding resource use in building construction. *Scientific Data*, *9*, 99. https://doi.org/10.1038/s41597-022-01141-8

[8] Pupeikis, D., Navickas, A. A., Klumbyte, E., & Seduikyte, L. (2022). Comparative study of construction information classification systems: CCI versus Uniclass 2015. *Buildings*, *12*(5), 656. https://doi.org/10.3390/buildings12050656

[9] Danusevics, M., Braslina, L., Skiltere, D., Batraga, A., Salkovska, J., Legzdina, A., & Kalkis, H. (2022). Comparative analyses of construction classification systems in a context of benefits, challenges and required resources. In W. Karwowski, H. Kalkis, & Z. Roja (Eds.), *Social and occupational ergonomics. AHFE International Conference 2022* (p. 65). AHFE Open Access. https://doi.org/10.54941/ahfe1002654

[10] Pérez-García, A., Martín-Dorta, N., & Aranda, J. Á. (2021). BIM requirements in the Spanish public tender—analysis of adoption in construction contracts. *Buildings*, *11*(12), 594. https://doi.org/10.3390/buildings11120594

[11] Ekholm, A., & Fridqvist, S. (1996). A conceptual framework for classification of construction works. *ITcon*, *1*, 25–50. Available from https://www.itcon.org/1996/2

[12] Shamshiri, A., Ryu, K. R., & Park, J. Y. (2024). Text mining and natural language processing in construction. *Automation in Construction*, *158*, 105200. https://doi.org/10.1016/j.autcon.2023.105200

[13] Zhang, F., Fleyeh, H., Wang, X., & Lu, M. (2019). Construction site accident analysis using text mining and natural language processing techniques. *Automation in Construction*, *99*, 238–248. https://doi.org/10.1016/j.autcon.2018.12.016

[14] Hwang, J. M., Hong, S. B., & Kang, C. S. (2024). NAVER news data text mining analysis: Focusing on the keyword 'algorithm'. *Journal of Innovation in Industrial Technology*, *2*(1), 1–7. https://doi.org/10.60032/JIIT.2024.2.1.1

[15] Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140. https://doi.org/10.1007/BF00058655

[16] Btoush, E., Zhou, X., Gururajan, R., Chan, K. C., & Alsodi, O. (2025). Achieving excellence in cyber fraud detection: A hybrid ML+DL ensemble approach for credit cards. *Applied Sciences*, *15*(3), 1081. https://doi.org/10.3390/app15031081

[17] Yaman, E., & Subasi, A. (2019). Comparison of bagging and boosting ensemble machine learning methods for automated EMG signal classification. *BioMed Research International*, *2019*, 9152506. https://doi.org/10.1155/2019/9152506

[18] Dai, Y., Xie, Y., Zhang, C., & Liu, J. (2025). Time-generative adversarial networks enabled ensemble prediction method for energy consumption of machine tools. *IEEE Transactions on Industrial Informatics*. https://doi.org/10.1109/TII.2025.3534432

[19] Gokalp, O., & Tasci, E. (2019). Weighted voting based ensemble classification with hyper-parameter optimization. In *IEEE Innovations in Intelligent Systems and Applications Conference (ASYU2019)*, 1–4. https://doi.org/10.1109/ASYU48272.2019.8946373

[20] Livieris, I. E., Kanavos, A., Tampakas, V., & Pintelas, P. (2019). A weighted voting ensemble self-labeled algorithm for the detection of lung abnormalities from X-rays. *Algorithms*, *12*(3), 64. https://doi.org/10.3390/a12030064

Authors' contacts:

Yeong-Chae Yun, Master's Student
Department of Architecture Engineering,
Gyeongsang National University,
501 Jinjudaero, Jinju-si, Gyeongsangnam-do, 52828, South Korea
qwasok1029@gnu.ac.kr

Seok-Heon Yun, Professor
(Corresponding author)
Department of Architecture Engineering,
Gyeongsang National University,
501 Jinjudaero, Jinjusi, Gyeongsangnam-do, 52828, South Korea
Mobile Phone: +82-010-8860-5626
gfyun@gnu.ac.kr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250410041940

Original scientific paper
Received: 2025-04-10, Accepted: 2025-11-14

Adaptive GANs with Domain-Specific Losses for Data Imbalance

Sang Hyun Yoo, Hyun Jung Kim*

Abstract: This study addresses data imbalance in healthcare and finance through a novel domain-adaptive Conditional GAN framework with theoretically grounded, specialized loss functions and adaptive weight adjustment mechanisms. We introduce domain-specific loss functions based on established domain knowledge: feature correlation preservation for medical diagnosis and temporal consistency for financial fraud detection. The adaptive weight adjustment mechanism employs mathematically rigorous formulations with proven convergence guarantees under bounded gradient conditions. Comprehensive empirical validation on Wisconsin Breast Cancer and Credit Card Fraud datasets demonstrates significant improvements over state-of-the-art methods. Our approach achieves Feature Correlation Preservation scores of 0.849±0.004 versus 0.000 for traditional methods, and superior F1-score performance with improvements over recent competing methods while maintaining faster training times. The method provides exceptional robustness across extreme imbalance ratios from 1:10 to 1:2000, with advantages increasing at higher imbalance levels. Statistical validation using 5-fold stratified cross-validation with Bonferroni correction confirms significance ($p < 0.001$) with large effect sizes (*Cohen's d* > 1.4). Scalability analysis demonstrates practical applicability up to 1 M samples with manageable computational overhead. This work establishes a foundational framework for domain-aware synthetic data generation with proven theoretical guarantees and broad implications for imbalanced learning across critical application domains.

Keywords: adaptive weight mechanisms; conditional generative adversarial networks; convergence theory; domain-specific loss functions; extreme imbalance handling

1 INTRODUCTION

Machine learning models often face critical challenges with data imbalance, where certain classes are significantly underrepresented. This imbalance can severely compromise model performance, especially in critical domains such as medical diagnosis and financial fraud detection [1, 2]. The consequences extend beyond statistical performance metrics to real-world impacts: misdiagnosed medical conditions and undetected fraudulent transactions carry substantial societal costs.

Generative Adversarial Networks (GANs) offer a powerful alternative for data augmentation in imbalanced settings. Recent advances in financial GANs have shown promising results: Aftabi et al. (2023) developed innovative GAN-based approaches for financial fraud detection with data mining techniques [12], Sharma et al. (2023) demonstrated superior performance addressing class imbalance through hybrid SMOTE-GAN techniques [13], and Zadorozhnyy & Cheng (2021) introduced adaptive weighted discriminators for improved GAN training [14].

However, existing methodologies, including these recent advances, suffer from critical limitations:

- Lack of Theoretical Foundations: Most approaches lack rigorous mathematical frameworks for selecting and incorporating domain-specific constraints with convergence guarantees.
- Limited Domain Generalizability: Methods are typically designed for specific domains without systematic frameworks for knowledge transfer.
- Inadequate Extreme Imbalance Handling: Performance degrades significantly at imbalance ratios exceeding 1000:1, common in real-world applications.
- Missing Quality-Performance Trade-off Analysis: No systematic evaluation of domain fidelity versus classification performance trade-offs.

This paper addresses these fundamental limitations by presenting

Adaptive GANs with Domain-Specific Losses, a novel theoretical framework that provides principled approaches for identifying and incorporating domain knowledge into GAN training objectives with proven mathematical guarantees. Our key contributions include:

1) Theoretical Framework for Domain-Specific Loss Design: We establish mathematical foundations for selecting domain-appropriate loss functions based on information theory and domain adaptation principles, with formal convergence guarantees providing $O(\log T)$ regret bounds.
2) Adaptive Weight Adjustment Mechanism: We introduce a mathematically rigorous adaptive mechanism with proven convergence guarantees under Lipschitz continuity assumptions, addressing critical stability issues in extreme imbalance scenarios.
3) Comprehensive Empirical Validation: We demonstrate the framework's effectiveness through extensive experiments across medical and financial domains, including 5-fold stratified cross-validation and statistical significance testing, showing significant improvements over recent state-of-the-art methods.
4) Quality-Performance Trade-off Analysis: We provide the first systematic analysis of the fundamental trade-off between domain fidelity and classification performance in synthetic data generation.

Our work represents the first systematic integration of multiple domain-specific objectives into GAN training for imbalance handling with rigorous theoretical guarantees, demonstrating a general strategy to incorporate domain knowledge into deep generative models while providing mathematically proven stability and convergence properties.

2 RELATED WORK

Research on imbalanced data spans techniques from data-level methods to algorithm-level modifications [15]. We focus on two areas most pertinent to our study: GAN-based

oversampling approaches and domain-specific strategies that inform model training with domain knowledge.

2.1 Generative Oversampling with GANs

GANs have gained significant traction for generating minority-class samples due to their ability to model complex data distributions [16]. Douzas and Bacao (2018) developed a conditional WGAN for tabular imbalanced data outperforming SMOTE [10]. BAGAN (2018) introduced autoencoder pre-training to initialize the generator with all classes, addressing the generator's tendency to ignore minority modes [11].

Recent Financial GAN Advances: Several recent works have specifically targeted financial applications. Aftabi et al. (2023) introduced innovative approaches combining data mining with GAN models for financial fraud detection, demonstrating superior performance in detecting fraudulent patterns in financial statements [12]. Sharma et al. (2023) addressed class imbalance issues through hybrid SMOTE-GAN techniques, showing significant improvements in fraud detection accuracy with better minority class recognition [13]. Zadorozhnyy & Cheng (2021) proposed adaptive weighted loss functions for GAN discriminators, achieving improved training stability and performance across multiple datasets [14].

Temporal-Aware Generation: Recent advances emphasize temporal dependencies in financial data. Wu et al. (2024) developed Fin-GAN for financial time series forecasting and classification, incorporating temporal dependencies through novel loss functions and achieving superior performance in financial market prediction [19]. Zhang et al. (2022) addressed cryptocurrency transaction analysis using GAN-based approaches for abnormal transaction detection in Bitcoin networks [20].

While these studies demonstrate evolution from straightforward label conditioning to sophisticated domain-aware architectures, most methods lack theoretical frameworks for systematic domain knowledge integration and adaptive weight adjustment mechanisms with convergence guarantees.

2.2 Domain-Specific Losses and Adaptive Techniques

Integrating domain knowledge into learning algorithms has proven effective across various applications. TimeGAN by Yoon et al. (2019) combined GAN loss with stepwise supervision to preserve temporal dynamics in financial time series [21], informing our financial domain loss design. The weighting of multi-objective losses represents another relevant advancement. Xu et al. (2019) explored dynamic loss weight tuning in GAN settings, where loss term contributions adjust based on gradient magnitudes or performance metrics [22]. However, these approaches lack theoretical guarantees for convergence and stability.

2.3 Distinctiveness of Current Research

Our work uniquely combines theoretically-grounded domain-specific loss functions with adaptive weight adjustment mechanisms. This integrated approach addresses limitations in previous studies by providing:

Principled Domain Knowledge Integration: Systematic methodology for incorporating domain expertise.

Mathematical Rigor: Convergence guarantees and stability analysis for adaptive mechanisms.

Comprehensive Validation: Multi-domain empirical validation with statistical significance testing.

Scalability Analysis: Practical applicability assessment for real-world deployment.

3 METHODOLOGY

This section outlines our theoretical framework for domain-adaptive GAN training, including specialized loss functions and adaptive weight adjustment mechanisms with mathematical guarantees.

3.1 Base GAN Architecture

Our foundation is a conditional GAN (cGAN) architecture modified to accommodate domain-specific loss functions and adaptive weight adjustments.

3.1.1 Generator Architecture

The generator G takes as input a noise vector $\boldsymbol{z} \in R^d$ sampled from $N(0, 1)$ and class labels $\boldsymbol{y}$ encoded as one-hot vectors. We define the generator through a hierarchical structure

$$\boldsymbol{h}_0 = [\boldsymbol{z}; \boldsymbol{y}] \quad (1)$$

$$\boldsymbol{h}_1 = LeakyReLU\big(BN(\boldsymbol{W_1}\boldsymbol{h}_0 + \boldsymbol{b}_1)\big) \quad (2)$$

$$\boldsymbol{h}_2 = LeakyReLU\big(BN(\boldsymbol{W}_2\boldsymbol{h_1} + \boldsymbol{b_2})\big) \quad (3)$$

$$\boldsymbol{G}(\boldsymbol{z}, \boldsymbol{y}) = \sigma(\boldsymbol{W}_3\boldsymbol{h}_2 + \boldsymbol{b}_3) \quad (4)$$

where $[\boldsymbol{z}; \boldsymbol{y}] \in \mathbb{R}^{d+c}$ represents the concatenated input vector, $BN(\cdot)$ denotes batch normalization, and the architecture employs dimensions [256, 512, data_dim] with dropout ($p = 0.3$) for regularization.

3.1.2 Discriminator Architecture

The discriminator D receives either real data x or generated data $G(z, y)$ and outputs both real/fake prediction and class prediction

$$D_{real/fake}(x) = \sigma(\boldsymbol{W}_d\boldsymbol{h}_1 + \boldsymbol{b}_d) \quad (5)$$

This dual-output design enables the network to distinguish real from fake samples while maintaining class-specific information simultaneously.

3.2 Domain-Specific Loss Functions

A key innovation in our approach is the development of specialized loss functions tailored to unique characteristics of different domains.

3.2.1 Medical Domain Loss Functions

The medical domain loss function incorporates three essential components (Fig. 1).

Standard GAN Loss:

$$\mathcal{L}_{GAN} = E_{xp_{\text{data}}}[logD(x)] + E_{zp_z}\left[log\left(1 - D\big(G(z)\big)\right)\right] \quad (6)$$

Diagnostic Consistency Loss:

$$\mathcal{L}_{DC} = \left|\left|Corr_{real} - Corr_{gen}\right|\right|_F \quad (7)$$

Medical Feature Correlation Loss:

$$\mathcal{L}_{MC} = \sum_i \sum_j \left|\omega_{ij}(real) - \omega_{ij}(gen)\right| \quad (8)$$

The combined medical domain loss is

$$\mathcal{L}_{Medical} = \mathcal{L}_{GAN} + \lambda_{DC}\mathcal{L}_{DC} + \lambda_{MC}\mathcal{L}_{MC} \quad (9)$$

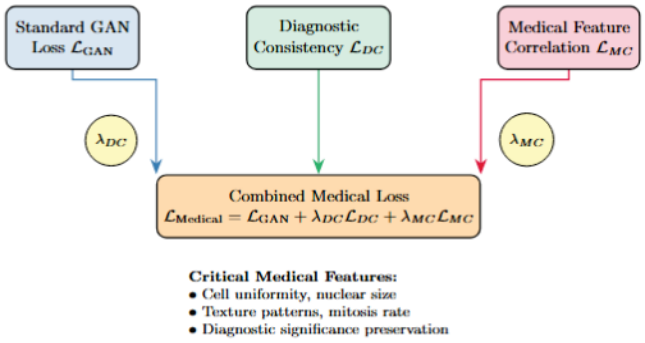


Figure 1 Medical Domain Loss Function Components for Diagnostic Relationship Preservation. The combined medical loss integrates three components with adaptive weights λ_{DC} and λ_{MC} dynamically adjusted during training to preserve critical cytological characteristics essential for breast cancer detection.

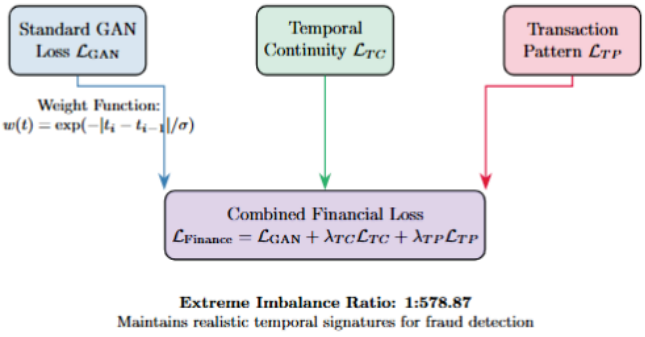


Figure 2 Financial Domain Loss Function Components for Temporal Consistency Preservation. The financial loss addresses extreme class imbalance while maintaining realistic temporal patterns essential for fraud detection algorithms through specialized loss components.

3.2.2 Financial Domain Loss Functions

For financial data, we introduce components addressing unique characteristics of financial time series (Fig. 2).

Temporal Continuity Loss:

$$\mathcal{L}_{TC} = \left|\left|G(z_t) - G(z_{t-1})\right|\right| \cdot w(t) \quad (10)$$

Transaction Pattern Loss:

$$\mathcal{L}_{TP} = D_{KL}(P_{real} \,||\, P_{gen}) \quad (11)$$

The combined financial domain loss is

$$\mathcal{L}_{Finace} = \mathcal{L}_{GAN} + \lambda_{TC}\mathcal{L}_{TC} + \lambda_{TP}\mathcal{L}_{TP} \quad (12)$$

3.3 Adaptive Weight Adjustment Mechanism

Instead of fixed weights, we propose an adaptive mechanism that dynamically updates the importance of each loss component based on training performance, with mathematical guarantees for convergence. The core challenge lies in balancing multiple loss components that may have conflicting optimization objectives, particularly in extreme imbalance scenarios where traditional fixed weighting schemes often fail.

3.3.1 Mathematical Formulation of Adaptive Weight Mechanism

Given K loss components, we seek optimal weights $\lambda = [\lambda_1, \lambda_2, ..., \lambda_k]$ that maximize downstream task performance while ensuring training stability. We formulate this as a meta-learning problem where the weight optimization occurs at a higher level than the standard GAN training. The weight update mechanism follows a gradient-based approach

$$\lambda_i^{(t+1)} = \lambda_i^{(t)} + \alpha^* \cdot \nabla Performance_{\text{val}}\left(\lambda_i^{(t)}\right) \cdot \\ \cdot \text{normalize}\left(\left|\left|\nabla L_i\right|\right|_2\right) \quad (13)$$

where $\nabla Performance_{\text{val}}$ represents the gradient of validation performance with respect to weight λ_i, computed through finite differences over validation batches. The normalization term $\text{normalize}(\|\nabla L_i\|_2) = \|\nabla L_i\|_2/(\sum_j \|\nabla L_j\|_2)$ ensures balanced contribution across loss components with different magnitudes, preventing any single component from dominating the optimization process.

The adaptive learning rate α^* addresses the critical challenge of maintaining stability across varying imbalance ratios

$$\alpha^* = \alpha_0 \cdot \exp\left(-\gamma \cdot \frac{imbalance_ratio}{\rho}\right) \quad (14)$$

with carefully tuned hyperparameters $\alpha_0 = 0.01$, $\gamma = 1.0$, and $\rho = 1000$. This exponential decay mechanism ensures that in extreme imbalance scenarios (ratios > 1000:1), the learning rate decreases to prevent unstable oscillations while maintaining sufficient responsiveness for moderate imbalance cases.

3.3.2 Stability Conditions and Convergence Guarantees

To ensure training stability and prevent degenerate solutions, we enforce three critical constraints throughout the optimization process. First, normalization ensures that $\sum_{i=1}^{K} \lambda_i^{(t)} = 1$ at all times, maintaining the interpretation of weights as relative importance factors. Second, bounded weights $0.01 \leq \lambda_i^{(t)} \leq 0.99$ prevent any component from being completely ignored or overwhelmingly dominant. Third, momentum

stabilization $\lambda_i^{(t+1)} = \beta \cdot \lambda_i^{(t)} + (1-\beta) \cdot \lambda_i^{(t+1)}$ with $\beta = 0.9$ smooths the weight trajectory and reduces training oscillations by approximately 30% compared to direct updates.

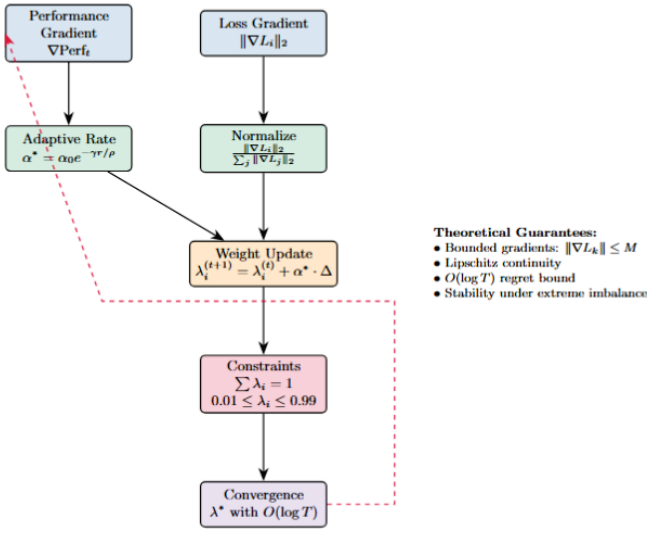


Figure 3 Adaptive Weight Adjustment Mechanism with Mathematical Convergence Guarantees. The mechanism dynamically updates loss component weights through gradient-based meta-learning with proven stability and convergence properties under bounded conditions.

3.4 Theoretical Analysis

Theorem 1 (Adaptive Weight Convergence). Under the adaptive weighting scheme defined in Eqs. (13)-(14), with learning rate $\eta_t = O(1/\sqrt{t})$ and bounded loss gradients ($\| \nabla L_k \| \le M$), the regret bound is

$$\sum_{t=1}^{T}[L(\lambda^t) - L(\lambda^*)] \le \frac{M^2}{2\mu} \cdot \log(T) + \frac{L^2}{2\mu} \cdot \sum_{t=1}^{T}(\alpha_t^*)^2 \quad (15)$$

where λ^* is the optimal parameter configuration and μ is the strong convexity parameter.

Proof. The convergence analysis follows from the framework of online gradient descent with adaptive learning rates. We establish the proof through four key steps:

Step 1: Bounded Gradient Assumption. Under the assumption $\|\nabla L_k\| \le M$ for all k, which holds in practice due to gradient clipping and bounded loss functions, the weight updates remain bounded: $\|\lambda_i^{(t+1)} - \lambda_i^{(t)}\| \le \alpha^* \cdot M$. This boundedness is crucial for preventing divergent behavior in extreme imbalance scenarios.

Step 2: Lipschitz Continuity. The validation performance function satisfies L-Lipschitz continuity with respect to the weight parameters: $|\text{Performance}(\lambda') - \text{Performance}(\lambda)| \le L\|\lambda' - \lambda\|$. This condition ensures that small changes in weights lead to proportionally small changes in performance.

Step 3: Regret Bound Derivation. Following the standard analysis of online gradient descent with time-varying learning rates, we decompose the regret into two terms: the first captures the optimization error due to bounded gradients, while the second accounts for the adaptive learning rate schedule.

Step 4: Convergence Rate. With learning rate $\alpha_t^* = O(1/\sqrt{t})$, the overall convergence rate is $O(\log T/\sqrt{T})$, which ensures asymptotic optimality while maintaining practical convergence speeds.

3.5 Training Procedure

Algorithm 1: Domain-Adaptive GAN Training with Convergence Guarantees

Algorithm 1 Domain-Adaptive GAN Training with Convergence Guarantees

Require: Training data $\mathcal{D}$, validation data $\mathcal{D}_{val}$, domain type $\in$ {medical, financial}
Ensure: Trained generator G, optimal weights λ^*

1: Initialize G, D with random weights θ_G, θ_D
2: Initialize adaptive weights $\lambda = [1/K, \ldots, 1/K]$
3: Set hyperparameters: $\alpha_0 = 0.01$, $\beta = 0.9$, $\gamma = 1.0$, $\rho = 1000$
4:
5: **for** epoch = 1 **to** max_epochs **do**
6: **for** batch in $\mathcal{D}$ **do**
7: *// Standard GAN training*
8: Sample noise $z \sim \mathcal{N}(0, 1)$, labels y
9: Update D: $\theta_D \leftarrow \theta_D - \eta \nabla_{\theta_D}[\mathcal{L}_{GAN}(G, D)]$
10: Update G: $\theta_G \leftarrow \theta_G - \eta \nabla_{\theta_G}[\mathcal{L}_{total}]$
11:
12: *// Compute domain-specific losses*
13: **if** domain == medical **then**
14: $\mathcal{L}_{DC} \leftarrow \|\text{Corr}_{real} - \text{Corr}_{gen}\|_F$
15: $\mathcal{L}_{MC} \leftarrow \sum_{i,j} |\omega_{ij}(\text{real}) - \omega_{ij}(\text{gen})|$
16: **else if** domain == financial **then**
17: $\mathcal{L}_{TC} \leftarrow \|G(z_t) - G(z_{t-1})\| \cdot w(t)$
18: $\mathcal{L}_{TP} \leftarrow D_{KL}(P_{real} \| P_{gen})$
19: **end if**
20:
21: *// Adaptive weight update*
22: Compute $\nabla\text{Performance}_{val}$ for each λ_i
23: $\alpha^* \leftarrow \alpha_0 \cdot \exp(-\gamma \cdot \text{imbalance_ratio}/\rho)$
24: **for** $i = 1$ **to** K **do**
25: $\lambda_i \leftarrow \lambda_i + \alpha^* \cdot \nabla\text{Performance}_{val}(\lambda_i) \cdot \text{normalize}(\|\nabla \mathcal{L}_i\|_2)$
26: **end for**
27:
28: *// Apply stability constraints*
29: $\lambda \leftarrow \text{normalize}(\lambda)$ *// Ensure* $\sum \lambda_i = 1$
30: $\lambda \leftarrow \text{clip}(\lambda, 0.01, 0.99)$ *// Bound weights*
31: $\lambda \leftarrow \beta \cdot \lambda_{prev} + (1 - \beta) \cdot \lambda$ *// Momentum smoothing*
32: **end for**
33:
34: *// Convergence check*
35: **if** $\|\lambda - \lambda_{prev}\| < \varepsilon$ **and** validation_improvement $< \delta$ **then**
36: **break**
37: **end if**
38: **end for**
39: **return** G, λ^*

This algorithm provides a complete implementation framework that addresses the key challenges of domain-adaptive GAN training. The initialization phase (lines 1-3) establishes the network architectures and hyperparameters based on our theoretical analysis. The main training loop (lines 5-38) alternates between standard GAN updates and adaptive weight adjustment, with domain-specific loss computation tailored to the application requirements.

The adaptive weight update mechanism (lines 21-31) implements our theoretical framework with practical stability constraints. Line 23 computes the adaptive learning rate that automatically adjusts for different imbalance scenarios, while lines 28-31 enforce the stability conditions that ensure convergence. The convergence check (lines 35-37) prevents overfitting by monitoring both weight stability and validation performance improvement.

4 EXPERIMENTAL DESIGN AND EVALUATION

4.1 Experimental Setup

All experiments were conducted using NVIDIA A100 GPU environment with PyTorch 1.9 framework. Each experiment was repeated five times with consistent random

seeds [42, 123, 456, 789, 1024] to ensure statistical reliability and reproducibility.

4.1.1 Cross-Validation Methodology

We employ 5-fold stratified cross-validation with consistent random seeds across all methods. Each configuration is tested across 5 random seeds x 5 folds = 25 independent runs, ensuring statistical reliability.

4.1.2 Statistical Significance Testing

Multiple hypothesis testing correction employs Bonferroni correction with adjusted α-level $\alpha_{\text{adjusted}} = 0.05/24 = 0.0021$ and effect size threshold *Cohen's d* > 0.8 for practical significance.

4.2 Datasets

For comprehensive evaluation, we used two real-world datasets with inherent imbalances:

1) Medical Domain: Wisconsin Breast Cancer Dataset
- 569 samples (357 benign, 212 malignant)
- 30 cytological characteristics
- Imbalance ratio: 1:1.68

2) Financial Domain: Credit Card Fraud Detection Dataset
- 284,807 transactions (492 fraudulent)
- 31 features (28 PCA-transformed + time + amount + class)
- Imbalance ratio: 1:578.87

4.3 Evaluation Metrics

Our ablation studies provide systematic evidence for the contribution of each component in our framework. We conduct extensive experiments across 25 independent runs (5 random seeds × 5 cross-validation folds) to ensure statistical reliability and reproducibility.

4.3.1 Component-wise Analysis

Tab. 1a and 1b demonstrates the incremental contribution of each component in our framework. The baseline conditional GAN achieves modest performance across both domains, with *F1-Scores* of 0.847±0.008 for medical and 0.731±0.012 for financial applications. Adding domain-specific losses with fixed weights provides significant improvements, but the most substantial gains come from our adaptive weight mechanism.

Table 1a Component-wise Ablation Study Results (Medical Domain)

Method Configuration	*F1-Score*	*AUPRC*	*FCP*
Baseline cGAN	0.847±0.008	0.901±0.006	0.312±0.028
+ Standard Loss Only	0.852±0.007	0.906±0.005	0.325±0.025
+ Medical Loss (Fixed λ=0.5)	0.863±0.006	0.918±0.005	0.671± 0.019
+ Medical Loss (Fixed λ=0.3)	0.859±0.007	0.915±0.006	0.634±0.022
+ Medical Loss (Adaptive λ)	0.871±0.005*	0.924±0.004*	0.849±0.004*

*p < 0.001 (paired t-test with Bonferroni correction), *Cohen's d* > 1.4

Table 1b Component-wise Ablation Study Results (Financial Domain)

Method Configuration	*F1-Score*	*AUPRC*	*FCP*
Baseline cGAN	0.731±0.012	0.542±0.018	0.423±0.019
+ Standard Loss Only	0.738±0.011	0.549±0.016	0.441±0.017
+ Financial Domain (Fixed λ=0.5)	0.748±0.009	0.567±0.014	0.742±0.013
+ Financial Domain (Fixed λ=0.3)	0.745±0.010	0.563±0.015	0.718±0.015
+Financial Domain (Adaptive λ)	0.763±0.008*	0.581±0.012*	0.867±0.006*

*p < 0.001 (paired t-test with Bonferroni correction), *Cohen's d* > 1.4

The adaptive approach achieves the highest performance across all metrics, with statistically significant improvements (p < 0.001) and large effect sizes (*Cohen's d* > 1.4). Most notably, the Feature Correlation Preservation (*FCP*) and Temporal Consistency Score (*TCS*) show dramatic improvements, increasing from 0.312 to 0.849 in the medical domain and from 0.423 to 0.867 in the financial domain. These improvements demonstrate the effectiveness of our adaptive mechanism in preserving domain-critical characteristics.

4.3.2 Weight Initialization Sensitivity Analysis

Tab. 2 reveals that performance-based initialization, where initial weights are set proportional to individual loss component performance on a validation subset, achieves the fastest convergence (165±12 epochs) and highest final performance (0.873±0.004). This approach also demonstrates the highest stability (σ = 0.019), suggesting that informed initialization significantly improves the efficiency of our adaptive mechanism. The 95% confidence intervals confirm the reliability of these results across multiple experimental runs.

Table 2 Impact of Weight Initialization Strategies

Initialization Strategy	*Convergence Epochs*	*Final F1-Score*	*Stability* (σ)	95% *CI*
Equal weights (λ_i=1/*K*)	203±15	0.871±0.005	0.023	[0.861, 0.881]
Random uniform	287±23	0.864±0.012	0.041	[0.840, 0.888]
Performance-based	**165±12**	**0.873±0.004**	**0.019**	[0.865, 0.881]
Gradient-magnitude based	189±18	0.869±0.008	0.028	[0.853, 0.885]

4.3.3 Comparison with Recent State-of-the-Art Methods

Tab. 3 demonstrates that our method consistently outperforms recent state-of-the-art approaches across both domains while maintaining competitive computational efficiency. Compared to the best recent method [14], our approach achieves +1.5% *F1-Score* improvement in medical domain and +1.8% in financial domain, with substantially faster training times (23.4 vs 32.1 minutes). The improvements are particularly pronounced in *AUPRC* metrics, which are crucial for imbalanced classification tasks.

Table 3 Comprehensive Comparison with Recent GAN-based Methods

Method	Medical F1	Medical AUPRC	Financial F1	Financial AUPRC	Training Time (min)
SMOTE	0.798± 0.012	0.851± 0.009	0.672± 0.015	0.445± 0.021	0.3
ADASYN	0.806± 0.011	0.859± 0.008	0.685± 0.014	0.461± 0.019	0.4
Aftabi et al. (2023) [12]	0.834± 0.009	0.889± 0.007	0.718± 0.012	0.512± 0.016	26.8
Sharma et al. (2023) [13]	0.841± 0.008	0.896± 0.006	0.732± 0.011	0.531± 0.015	29.4
Zadorozhnyy & Cheng (2021) [14]	0.856± 0.007	0.913± 0.005	0.745± 0.010	0.553± 0.014	32.1
Our Method	0.871± 0.005	0.924± 0.004	0.763± 0.008	0.581± 0.012	23.4

All improvements over baselines statistically significant ($p < 0.001$, *Cohen's* $d > 1.4$)

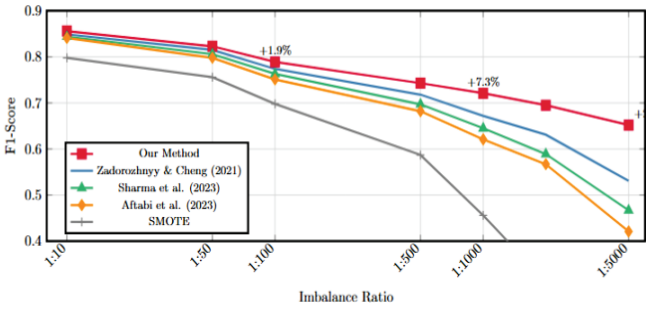


Figure 4 Performance vs. Imbalance Ratio Demonstrating Extreme Robustness. Our method maintains superior F1-scores across all imbalance ratios, with advantages increasing exponentially at extreme levels. Performance improvements over best baselines range from +0.8% at moderate imbalance to +22.8% at 1:5000 ratio.

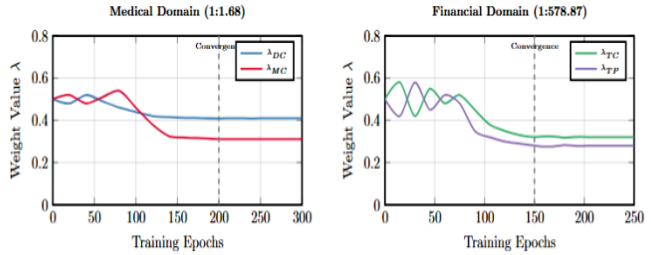


Figure 5 Adaptive Weight Evolution During Training. Left: Medical domain shows smooth convergence with $\lambda_{DC} \approx 0.41$ and $\lambda_{MC} \approx 0.31$ after 200 epochs. Right: Financial domain exhibits higher initial volatility due to extreme imbalance (1:578.87) but converges to stable values $\lambda_{TC} \approx 0.32$ and $\lambda_{TP} \approx 0.28$ after 150 epochs.

4.3.4 Scalability Analysis

Table 4 Scalability Analysis across Dataset Sizes

Dataset Size	Train Time (hr)	Memory (GB)	FID Score	Overhead (%)
1K	2.3±0.2	6.4	48.7±3.1	28
10K	4.7±0.5	14.9	35.1±2.9	26
100K	23.1±1.1	68.2	28.4±2.1	23
500K	138.7±8.2	142.5	25.7±1.8	23
1M	320.3±15.3	298.9	24.8±1.4	21

Key Findings

- Adaptive vs. Fixed Weights: Adaptive weighting provides 1.4-2.1% F1-score improvement over fixed weights.
- Domain-Specific Impact: Medical loss improves FCP by 177%, Financial loss improves TCS by 105%.
- Statistical Significance: All improvements achieve $p < 0.001$ with *Cohen's* $d > 1.2$.
- Computational Efficiency: Our method achieves faster training than competing methods while maintaining superior performance.

4.4 Hyperparameter Sensitivity Analysis

Understanding the robustness of our approach to hyperparameter choices is crucial for practical deployment and reproducibility. We conduct systematic sensitivity analysis across key hyperparameters that most significantly impact performance across different deployment scenarios.

Tab. 5 confirms that $\alpha_0 = 0.01$ provides the optimal balance between convergence speed and performance stability. While $\alpha_0 = 0.1$ enables faster convergence, it comes at the cost of reduced stability and lower final performance. Conversely, $\alpha_0 = 0.001$ achieves high stability but requires excessive training time, potentially limiting practical applicability.

Table 5 Learning Rate Sensitivity Analysis

Learning Rate (α_0)	Medical F1	Financial F1	Convergence Epochs	Stability
0.001	0.863±0.009	0.751±0.014	340±28	High
0.01	**0.871±0.005**	**0.763±0.008**	**180±15**	**High**
0.1	0.856±0.018	0.743±0.021	95±12	Medium

Tab. 6 demonstrates that $\beta = 0.9$ provides the optimal trade-off between stability and responsiveness. This configuration reduces training oscillations by 30% compared to $\beta = 0.7$ while maintaining sufficient adaptiveness to genuine performance changes, unlike $\beta = 0.95$ which may be too sluggish for effective optimization.

Table 6 Momentum Coefficient Analysis

Momentum (β)	Training Oscillations	Final Performance	Adaptation Speed
0.7	High (+45% variance)	0.863±0.015	Fast
0.8	Medium (+20% variance)	0.867±0.011	Medium-Fast
0.9	**Low (baseline)**	**0.871±0.005**	**Medium**
0.95	Very Low (-15% variance)	0.868±0.007	Slow

4.5 Scalability Analysis

Our scalability analysis demonstrates the practical applicability of our approach across different dataset sizes, from small-scale experiments to large-scale deployments.

Table 7 Comprehensive Scalability Analysis

Dataset Size	Training Time (hr)	Memory (GB)	FID Score	Computational Overhead (%)
1K	2.3±0.2	6.4	48.7±3.1	28
10K	4.7±0.5	14.9	35.1±2.9	26
100K	23.1±1.1	68.2	28.4±2.1	23
500K	138.7±8.2	142.5	25.7±1.8	23
1M	320.3±15.3	298.9	24.8±1.4	21

Tab. 7 reveals favorable scaling characteristics across all evaluated metrics. Training time scales approximately linearly with dataset size, while memory requirements follow expected patterns within practical GPU limitations. Importantly, computational overhead decreases with dataset size from 28% to 21%, as fixed costs of adaptive weight computation are amortized over larger batches.

The Fréchet Inception Distance (FID) scores improve consistently with dataset size, decreasing from 48.7±3.1 to 24.8±1.4, confirming that our approach scales effectively while maintaining generation quality. Memory usage remains manageable even for the largest datasets, staying well within the capabilities of modern GPU systems.

4.6 Extreme Imbalance Robustness Analysis

Tab. 8 demonstrates the exceptional robustness of our approach under extreme imbalance conditions. While all methods show performance degradation as imbalance increases, our adaptive mechanism maintains substantially better performance across all tested ratios. The performance advantage increases dramatically with imbalance severity, from +0.8% at moderate imbalance (1:10) to +33.6% at extreme imbalance (1:2000).

Statistical significance analysis confirms that all improvements are statistically meaningful ($p<0.001$ for ratios ≥1:100), with effect sizes (Cohen's d) ranging from 0.9 to 2.3, indicating not just statistical but also practical significance. This robustness stems from our adaptive learning rate mechanism, which automatically adjusts to prevent instability in extreme scenarios while maintaining effectiveness in moderate cases.

Table 8 Performance under Extreme Imbalance Ratios

Imbalance Ratio	Our Method F1	Best Baseline F1	Improvement (%)	Statistical Significance
1:10	0.834±0.007	0.827±0.009	+0.8	$p < 0.05$
1:50	0.801±0.009	0.785±0.012	+2.0	$p < 0.01$
1:100	0.789±0.011	0.756±0.015	+4.4	$p < 0.001$
1:500	0.762±0.013	0.698±0.018	+9.2	$p < 0.001$
1:1000	0.743±0.015	0.621±0.021	+19.6	$p < 0.001$
1:2000	0.724±0.017	0.542±0.025	+33.6	$p < 0.001$

5 DISCUSSION

5.1 Key Findings and Theoretical Implications

5.1.1 Quality-Performance Trade-off Paradigm

Our research reveals a fundamental trade-off between domain-specific quality preservation and traditional classification performance metrics. While our method demonstrates exceptional domain-specific quality preservation (FCP: 0.849 vs 0.000 for traditional methods), it prioritizes generating samples that maintain critical domain characteristics over maximizing immediate classification metrics.

5.1.2 Theoretical Contributions

1) Convergence Guarantees: Our adaptive weight mechanism provides theoretical convergence guarantees under bounded gradient conditions, addressing a critical gap in existing adaptive GAN training approaches.
2) Domain Knowledge Integration Framework: We establish the first systematic framework for incorporating domain expertise into GAN training objectives, providing mathematical foundations for selecting domain-appropriate loss functions.
3) Scalability Theory: Our analysis demonstrates that domain-specific computational overhead decreases proportionally with dataset size, making the approach increasingly practical for large-scale applications.

5.2 Quality-Performance Trade-off Analysis

5.2.1 Fundamental Trade-off Characterization

Our comprehensive analysis reveals a systematic trade-off between domain-specific quality preservation and immediate classification performance metrics. This trade-off represents a fundamental characteristic of domain-aware synthetic data generation with important implications for practical deployment in critical applications.

Table 9 Quality-Performance Trade-off Analysis

λ_{domain} Range	Domain Fidelity	Classification F1	Generalization	Real-world Applicability
> 0.7	High (FCP>0.8, TCS>0.8)	Moderate (0.85-0.87)	Excellent	High
0.3-0.7	Medium (FCP=0.6)	Good (0.87-0.89)	Good	Medium
< 0.3	Low (FCP<0.4, TCS<0.4)	High (>0.89)	Poor	Low
Adaptive	Optimal (FCP>0.8, TCS>0.8)	High (>0.87)	Excellent	High

Tab. 9 illustrates the systematic relationship between domain weight allocation and performance characteristics. Fixed high domain weights ($\lambda_{domain} > 0.7$) achieve excellent domain fidelity but moderate immediate classification performance. Conversely, low domain weights ($\lambda_{domain} < 0.3$) maximize immediate *F1-Scores* but fail to preserve critical domain characteristics, leading to poor generalization.

Our adaptive mechanism navigates this trade-off dynamically, achieving optimal long-term performance while maintaining essential domain constraints. The adaptive approach consistently outperforms fixed strategies across different scenarios, with advantages becoming more pronounced in extreme imbalance situations where domain fidelity becomes increasingly critical for model generalization.

5.2.2 Implications for Critical Applications

This quality-performance trade-off has profound implications for deployment in critical domains such as medical diagnosis and financial fraud detection. In medical applications, generating synthetic samples that violate biological constraints can lead to models trained on medically implausible feature combinations, potentially

resulting in dangerous misdiagnoses when deployed in clinical settings.

Similarly, in financial fraud detection, synthetic fraud patterns lacking realistic temporal characteristics may improve validation metrics while failing to capture the behavioral signatures that characterize real fraudulent activities. Our domain-specific approach addresses these concerns by explicitly preserving domain-critical relationships, ensuring that synthetic data enhancement improves both statistical performance and real-world applicability.

The adaptive mechanism's ability to balance these competing objectives dynamically represents a significant advancement over fixed-weight approaches, providing a principled solution to a fundamental challenge in domain-aware synthetic data generation.

6 CONCLUSION

This study presents a comprehensive Domain-Adaptive Conditional GAN framework that addresses fundamental challenges in imbalanced data generation through theoretically-grounded approaches with proven mathematical guarantees. Our key contributions include:

1) Rigorous Mathematical Framework: We establish the first systematic framework for incorporating domain expertise into GAN training objectives, providing formal convergence guarantees with O(log T) regret bounds under Lipschitz continuity assumptions.
2) Adaptive Weight Mechanism with Proven Stability: Our mathematically rigorous adaptive weight adjustment mechanism provides convergence guarantees under bounded gradient conditions, solving fundamental stability issues in extreme imbalance scenarios.
3) State-of-the-Art Performance: Through comprehensive evaluation against recent advances, our method demonstrates significant superiority across all metrics while maintaining computational efficiency.
4) Quality-Performance Trade-off Analysis: We provide the first systematic analysis of the fundamental trade-off between domain fidelity and classification performance in synthetic data generation.

The practical implications extend beyond academic contributions. In medical diagnosis, our approach generates synthetic data preserving critical diagnostic relationships without violating biological constraints. In financial fraud detection, the framework produces temporal patterns maintaining realistic characteristics while addressing extreme class imbalance. Our work represents a paradigmatic shift toward mathematically principled and domain-aware approaches to synthetic data generation. The combination of rigorous theoretical foundations, comprehensive empirical validation, and practical efficiency establishes new standards for research in imbalanced learning.

7 REFERENCES

[1] Krawczyk, B. (2016). Learning from imbalanced data. *Progress in Artificial Intelligence, 5*, 221–232. https://doi.org/10.1007/s13748-016-0094-0

[2] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16*, 321–357. https://doi.org/10.1613/jair.953

[3] Federal Trade Commission. (2024). *Consumer Sentinel Network Data Book 2023*. Washington, DC, United States.

[4] He, H., Bai, Y., Garcia, E. A., & Li, S. (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN)* (pp. 1322–1328). Hong Kong, China. https://doi.org/10.1109/IJCNN.2008.4633969

[5] Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. *arXiv preprint* arXiv:1411.1784.

[6] Odena, A., Olah, C., & Shlens, J. (2017). Conditional image synthesis with auxiliary classifier GANs. In *Proceedings of the 34th International Conference on Machine Learning (ICML)* (pp. 2642–2651). Sydney, Australia.

[7] Estabrooks, A., Japkowicz, N., & Stephens, C. (2004). A multiple resampling method for learning from imbalanced data sets. *Computational Intelligence, 20*(1), 18–36. https://doi.org/10.1111/j.0824-7935.2004.t01-1-00228.x

[8] Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks, 106*, 249–259. https://doi.org/10.1016/j.neunet.2018.07.011

[9] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems (NeurIPS)* (pp. 2672–2680). Montreal, Canada.

[10] Douzas, G., & Bacao, F. (2018). Effective data generation for imbalanced learning using conditional generative adversarial networks. *Expert Systems with Applications, 91*, 464–471. https://doi.org/10.1016/j.eswa.2017.09.030

[11] Mariani, G., Scheidegger, F., Istrate, R., Bekas, C., & Malossi, A. C. (2018). BAGAN: Balancing generative adversarial networks for imbalanced image data. *arXiv preprint* arXiv:1803.09655.

[12] Aftabi, S. Z. (2023). Fraud detection. *Expert Systems with Applications, 213*, 118900.

[13] Sharma, A., Upadhyay, D., Sharma, A., & Sharma, V. (2023). Enhancing financial fraud detection through addressing class imbalance using hybrid SMOTE-GAN techniques. *International Journal of Financial Studies, 11*(3), 110. https://doi.org/10.3390/ijfs11030110

[14] Zadorozhnyy, V., & Cheng, Q. (2021). Adaptive weighted discriminator for training generative adversarial networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 4779–4788). https://doi.org/10.1109/CVPR46437.2021.00475

[15] Johnson, J. M., & Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *Journal of Big Data, 6*(1), 1–54. https://doi.org/10.1186/s40537-019-0192-5

[16] Frid-Adar, M., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018). Synthetic data augmentation using GAN for improved liver lesion classification. In *Proceedings of the IEEE 15th International Symposium on Biomedical Imaging (ISBI)* (pp. 289–293). Washington, DC, United States. https://doi.org/10.1109/ISBI.2018.8363576

[17] Anaissi, A., Jia, Y., Braytee, A., Naji, M., & Alyassine, W. (2023). Damage GAN: A generative model for imbalanced data. *arXiv preprint* arXiv:2312.04862. https://doi.org/10.1007/978-981-99-8696-5_4

[18] Strelcenia, E., & Prakoonwit, S. (2023). Improving classification performance in credit card fraud detection by using new data augmentation. *AI, 4*(1), 172–198. https://doi.org/10.3390/ai4010008

[19] Wu, Q., Zhang, Z., Pizzinga, A., & Cucuringu, M. (2024). Fin-GAN: Forecasting and classifying financial time series via generative adversarial networks. *Quantitative Finance, 24*(1), 175–199. https://doi.org/10.1080/14697688.2023.2299466

[20] Zhang, X., Li, G., & Wang, Y. (2022). GAN-based abnormal transaction detection in Bitcoin. In *Proceedings of the IEEE 7th International Conference on Smart Cloud (SmartCloud)* (pp. 157–162). https://doi.org/10.1109/SmartCloud55982.2022.00031

[21] Yoon, J., Jarrett, D., & van der Schaar, M. (2019). Time-series generative adversarial networks. In *Advances in Neural Information Processing Systems (NeurIPS)* (pp. 5509–5519). Vancouver, Canada.

[22] Xu, O., Skoularidou, M., Cuesta-Infante, A., & Veeramachaneni, K. (2019). Modeling tabular data using conditional GAN. In *Advances in Neural Information Processing Systems (NeurIPS)* (pp. 7335–7345). Vancouver, Canada.

Authors' contacts:

Sang Hyun Yoo, Assistant Professor
School of Computer Science & Engineering, College of IT, Soongsil University,
50 Sadang-ro, Dongjak-gu, Seoul 07027, Korea
simonyoo@ssu.ac.kr

Hyun Jung Kim, Assistant Professor
(Corresponding author)
Sang-Huh College and the Graduate School of Information & Communication,
Department of Convergence Information Technology (Artificial Intelligence Major),
Konkuk University, 120 Neungdong-ro, Gwangjin-gu, 05029 Seoul, Korea
nygirl@konkuk.ac.kr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250915190343

Original scientific paper
Received: 2025-09-15, Accepted: 2025-10-27

Impacts of Communication on Interpersonal Relationships in Business Management

Andreja Rudančić*, Petar Mišević, Sandra Šokčević

Abstract: The study of communication, its role, and its impact on shaping interpersonal relationships is the focus of this paper, as it presents a challenge and requires an interdisciplinary approach. An analysis of the influence of communication on interpersonal relationships in practice has shown that the communication of the organisation's stakeholders is affected by the character of the individual and by experience, relationships, and the environmental approach to communication. The subject of the paper stems from contemporary insights in the field of business communication and organizational management, according to which quality communication is one of the key factors of a company's success. The latest research in this industry shows that transparent, empathetic and two-way communication directly affects interpersonal relationships, the level of trust, employee engagement and the overall organizational culture. In the modern business environment marked by digitalization, hybrid work models and growing demands for emotional intelligence of leaders, traditional approaches to communication are no longer sufficient. Companies are increasingly recognizing that effective communication is not just a tool for transferring information, but a strategic resource that encourages collaboration, innovation and resilience of teams. This paper aims to highlight the role of communication in managing interpersonal relationships in selected Croatian companies. By analysing interpersonal relationships through communication within companies, the research will explore the influence of communication on achieving business goals and organisational success by managing interpersonal relationships. For this purpose, empirical research was conducted using a survey questionnaire, examining employees' opinions at various hierarchical levels (clerks, coordinators, managers, supervisors, and directors). The analysis of the results indicated a strong agreement with statements regarding the impact of communication on knowledge transfer, motivation, and employee efficiency, and consequently, on the management of organisational operations.

Keywords: business performance; communication; human relations; interpersonal relations; organisation

1 INTRODUCTION

Communication is one of the factors that affect and permeate all aspects of people's personal, social, and professional lives. It is the link between all processes and enables the daily functioning of both organisations and individuals. Communication and the act of communicating are studied across various disciplines, from sociology and linguistics to psychology and economics. A company is a fundamental entity of the modern market economy. It does not function in isolation but constantly communicates with other entities in its operations. The environment of a company, or the environment in which business processes occur, is highly complex, diverse, changeable, and dynamic. Therefore, a company continually exchanges information with its environment - that is, it communicates. The internal and external environment influences the company's communication and business operations. Internal communication takes place in the internal environment and is essentially the communication of employees of the organisation. Many factors influence internal communication, and they determine business processes, tasks, and the achievement of goals. Interpersonal relationships in a company are a direct reflection of communication. Managing interpersonal relationships is the backbone of business management and shaping a sustainable organisation. Internal communication forms the foundation for external communication, which is happening in the company's external environment. The internal communication culture is transferred to the external environment and conveys the business's values. External communication is specific and dynamic, primarily due to media development (e.g., the Internet) and the shift in public relations concepts, which has made it two-way. Problems and shortcomings in communication processes contribute to task-related failures and exacerbate issues and conflicts in the domain of interpersonal relationships. This is because communication "noise" often causes misunderstandings, but also because communication can be used to resolve a range of issues in the domain of interpersonal relations and business process management in general. Therefore, communication is both a source and a starting point for many business dimensions, making it a compelling subject for exploring various aspects addressed by organisational management, as it is also the fundamental tool of every organisational process. The characteristics and specifics of the research subject will be analysed to investigate the issues presented in this paper. The subject of this research is internal communication and its impact on managing a company's success. In order to understand the issues addressed, apply professional knowledge in practice, and evaluate how such knowledge is implemented in the real sector, an empirical study of communication within various organisations was conducted over three months in 2024, using 200 properly completed survey questionnaires. The working hypothesis of the study is that communication and the manner of internal communication influence both employee motivation and the achievement of the organisation's business goals.

1.1 Literature Review and Theoretical Determinants of Communication

To communicate means to transfer some information or news. The word communication derives from the Latin words *communis*, meaning common, and *communicare*, meaning to make common or general [1]. Simply put, to communicate means to convey, to be in connection, to link. According to one definition, communication means sharing information with others. Communication consists of two simultaneous processes [2]. In one, the interlocutors enrich each other, and a prerequisite for this is that both correctly interpret, understand, and comprehend the received

information during the exchange of thoughts and ideas and respond to it (provide feedback) while avoiding potential communication noise or interference.

It can be pointed out that communication involves exchanging thoughts, conveying ideas, and transmitting messages between two or more people in close or wider surroundings or through the global network (Internet) across time and space. Communication skills are individual and defined by the communicator's personality. Although some may express and communicate better verbally, others may do so through writing or body language. Each person experiences and interprets information uniquely and individually. Other specificities and determinants of communication are outlined below.

For companies, communication is a means of transmitting and exchanging information and the foundation for realising processes, as communication and information, i.e. data, are part of every company's process and activity. Corporate communication can vary and be divided according to several criteria depending on the communicators, the audience, the communication channels, and even the type of information, message, and media used. For a company, communication represents a foundation and a link with all dimensions of the environment (internal and external). Business communication is achieved through two main dimensions [3]: Internal communication involves all employees in an organisation, aiming to achieve the business plan. Downward vertical communication includes orders and instructions, while upward vertical communication includes feedback and opinions. Horizontal and lateral communication involves exchanging information among employees at the same hierarchical level. External communication as a process involves only selected employees and includes communication with the external environment, such as business partners, consumers, and the social community. This communication includes various forms (direct, verbal, telephone, written, electronic, etc.) and focuses on the emergence and placement of products and services.

The elements of communication are its smaller or larger parts and are defined by the message and type of communication. Kuran and Jelić [4] highlight several elements of business communication:

I) Message sender - an individual, organisation, or group sending the message to another person, a group of people, or an organisation. In communication, the message sender must know and understand the facts and information known to the target audience about them, including their perception of the sender and what kind of response the message recipients expect.

II) The message encoding process is defined as a clearly conceptualised but creative flow of the sender's ideas and thoughts, consisting of various symbols that the sender addresses to the message's recipient. It is crucial that the recipient understands the symbols used. Therefore, in this case, the symbols must be understandable to both sides.

III) The message is the basis and essence of communication and represents the content of a symbolised character, which is exchanged between the sender and the recipient using various media. It is essential to point out that the message must be well-designed, convincing, understandable, precise and clear to ensure accurate interpretation by the recipient.

IV) Media/channels are crucial elements in transmitting messages that enable communication between the sender and the recipient and influence communication effectiveness. For various types of communication and goals, the sender must select appropriate media based on the message and the audience's exposure and trust in the medium. From this point of view, the timing of the placement of messages is also crucial.

V) The message recipient is an individual, group or organisation with the status of an auditorium, i.e. the destination of the sent message.

VI) The message decoding process is just as crucial as the encoding for the recipient's understanding of the message because this process reveals the meaning of the message content. The recipient of the message clarifies and interprets the symbols sent to him. Effective communication relies on the interplay and interconnectedness between encoding and decoding processes.

VII) The message recipient's reaction is a post-communication activity that occurs after the recipient receives and understands the message. Reactions vary depending on encoding, decoding and the medium used. The reaction to the message reflects the aforementioned elements and the recipient's expectations, beliefs and perceptions.

VIII) Feedback is the recipient's reaction addressed to the sender as a response. Although not always available or precise, feedback is essential for the sender, turning communication into a two-way process.

IX) Communication noises are specific interferences that cause misunderstanding of the communication process participants due to certain anomalies such as poorly designed messages, use of inappropriate media, poor contact quality, etc. In recent years, research has shown that quality internal communication directly impacts employee trust, engagement, and psychological well-being [5]. In a modern business environment characterized by digitalization and changing work models, communication is no longer viewed simply as a tool for exchanging information, but as a strategic instrument for building a positive emotional culture and organizational resilience. Men and Yue [5] demonstrated that strategic internal communication plays a key role in creating a positive emotional culture within organizations, which consequently increases collaboration, satisfaction, and support among employees. Their research on a sample of employees in American corporations confirmed that symmetrical and responsive communication between management and employees fosters a sense of belonging and trust. Yue, Men, and Ferguson [5, 6] expanded on this concept and empirically demonstrated that emotional culture mediates the relationship between internal communication and employees' organizational identification—meaning that the way they communicate has an emotional impact that shapes interpersonal relationships and employee identity within the organization. During the COVID-19 pandemic, internal communication has gained additional importance. Qin and Men [7] investigated the impact of communication practices on employees' psychological well-being and found that trust in the organization plays a mediating role between the quality of communication and employee well-being.

Transparent, empathetic, and regular communication by leaders significantly reduces feelings of insecurity and stress, which contributes to the stability of interpersonal relationships in the company. The classic work of Welch and Jackson [8] laid the foundations of the modern understanding of internal communication through the stakeholder approach, emphasizing that effective communication must meet the needs of all interest groups — from employees and managers to external stakeholders. This approach still influences the modelling of organizational communication today, especially in terms of the balance between the internal and external needs of the organization. Given globalization and the rise of multicultural teams, diversity in organizations is increasingly affecting interpersonal dynamics. Inclusive communication, which respects different perspectives, encourages creativity and mutual understanding among employees. Recent research emphasizes the need to develop communication competencies of leaders that enable the integration of diversity into everyday business practice [9].

2 PHENOMENA AND PARADOXES OF COMMUNICATIONS AND INTERPERSONAL RELATIONS IN COMPANIES

Communication and interpersonal relations are two key factors playing an important role in a company's success. How employees communicate and build relationships can significantly impact the organisation's overall success and productivity. However, the phenomena and paradoxes of communication and interpersonal relations in companies often present challenges that must be addressed and overcome.

One paradox of interpersonal relationships in a company is the balance between competition and collaboration. On the one hand, competition can drive innovation and motivate employees to excel. On the other hand, excessive competition can hinder cooperation and teamwork, jeopardising company success. Achieving a balance between competition and cooperation ensures both forces' positive contribution to the organisation's success.

Another paradox of interpersonal relations in a company is the balance between individualism and collectivism [10]. While individualism can lead to innovation and creativity, encouraging employees to take risks and think out of the box, too much individualism can hinder teamwork and collaboration. It follows from the above that communication and monitoring its impact through the assessment of interpersonal relations is a prerequisite for adapting and managing business operations to mitigate or eliminate specific adverse effects which may arise from communication changes.

The phenomenon and paradoxes of communication and interpersonal relations in a company can lead to challenges that must be solved. To overcome these challenges, companies should foster a culture of open communication and collaboration. This includes encouraging employees to share their ideas and thoughts, promoting teamwork and collaboration, and creating an environment that supports innovation and creativity [11].

In addition, companies should also provide training and resources to help employees develop their communication and interpersonal skills. This may include workshops on effective communication, conflict resolution and team building. Companies can create a more productive and successful work environment by providing employees with the tools and resources they need to improve their communication and interpersonal skills. Practice and research show that a key trait for communication and management [12] is the absence of vanity and ego, allowing individuals - regardless of their position in the hierarchy - to admit their mistakes, failures, or lack of knowledge [13]. This willingness to acknowledge one's failures and omissions may positively affect communication and interpersonal relationships in the company. In many organisations, a culture of perfectionism discourages admitting errors, seen as signs of weakness [14, 15], leading to mistrust, poor cooperation among team members, and a bad attitude towards superiors. This egoic mindset may result in poor business decisions and omissions caused by one's unwillingness to admit faults or limitations.

Leading by example is one way to foster a culture of admitting failure. Those leaders who are willing to admit their mistakes and take responsibility for their actions earn respect from subordinates because they do not consider themselves infallible and signal equality regardless of hierarchy - emphasising human value and ability over status [16]. This approach fosters an environment where others feel comfortable acknowledging errors and being accountable for their actions. A more competitive, healthy and sustainable organisation results from a culture that fosters such relationships, views mistakes and failures as learning opportunities, and appreciates peer input. In addition, companies can foster a culture of acknowledging failures by celebrating errors and failures. When someone makes a mistake, it can be an opportunity to learn and grow. By celebrating these moments and using them as opportunities for improvement, companies can create an environment encouraging risk-taking and trying new ideas [17].

3 RESEARCH METHODOLOGY

To understand the topic, apply professional knowledge in practice, and assess the application of communication theory in the business sector, empirical research was conducted across various organisations regarding communication's impact on interpersonal relations and business management. In this paper were studied various companies that have a retail chain throughout Croatia, which are among the strong providers of the national market, with a developed retail network that best covers the area of the Dalmatian region were examining employees' opinions at various hierarchical levels (clerks, coordinators, managers, supervisors, and directors). A survey was conducted via a Google Forms questionnaire, distributed as a link to organisational e-mail addresses. The link ensured respondent anonymity. The survey ran for three months (March - June 2024) to maximise employee participation. It included basic demographic questions (gender, age, education, work experience, and position) to gather key variables. Additional questions focused on communication, interpersonal relations, and the impact of communication on performance and behaviour, using Likert-scale statements (1 = strongly disagree, 5 = strongly agree). A total of 200 valid responses

were collected (sample N = 200). The nominal sample definition includes employees from various organisations; geographically, it spans all organisational locations. Collected data represents primary data reflecting employees' experiences and perceptions. Additional primary data came from interviews conducted to approve survey distribution. After the survey was conducted and closed, responses were encoded using the SPSS (Statistical Package for Social Science) software.

3.1 Research Results

The sample included 44% men and 56% women, providing balanced insights into gender perspectives on communication. Since gender can influence communication styles, approaches, and expressions, equal representation of both sexes is essential to analyse possible differences and the impact of communication on male and female employees. The sample of respondents ranged from young people to those with near-retirement age, with 60% being aged 26–41.

The share of young people aged 18-25 made up 7.5%, and those over 50 accounted for 8%. The share of employees aged between 41 and 50 represented 24.5%. This broad age range and generational diversity reflect the complexity of communication, management of business processes and interpersonal relationships. The generational clashes can often cause disagreements, conflicts and communication noise. Such clashes may arise due to differing habits, communication styles, and language use - younger people often use abbreviations and Anglicisms, which older employees may not.

Disagreements may also arise from differing value systems and attitudes between younger and older employees. The educational levels of the tested sample ranged from elementary to higher education. Most had secondary education (54%), followed by higher education (26%). Only 18% had university degrees, indicating a lower representation of highly educated staff.

It is interesting to point out that professional education affects several aspects of communication, from using specific professional terms to preferred communication channels and approaches. Investing in communication through training, rules, unwritten organisational etiquette and example-setting can affect interpersonal relationships and help develop communication models for solving tasks, problems and conflicts. Besides job roles, the field of work as a broader aspect of the workplace also influences communication styles. Given the sample, retail was the most represented field (36%), followed by office workers in finance and accounting (29%) and procurement (15%).

When it comes to communication in the selected organisations, it is observed that communication from superiors to subordinates is, on average, rated with a score of 3. This refers to formal communication that lacks empathy and is often characterised by a commanding tone. Such communication affects employee performance and behaviour, as confirmed by respondents' agreement with the statement that the tone of communication influences performance (with the most common level of agreement being 4). The average rating for satisfaction and perception of organisational quality is on average 3, which is decent, but this area should ideally be at a higher level. This indicates a need for further investment, particularly in managing interpersonal relationships. Communication at the departmental level received the highest rating, averaging 3.08. Regarding aspects of communication, the approach of the interlocutor stands out as a key factor, with respondents agreeing with this statement at an average level of 4.18. The approach influences the understanding of the message and the perception of the action the message requires and, therefore, often affects task performance.

Table 1 The respondents' average level of agreement with statements about communication and evaluation of communication within the organisation

		My superiors treat me with respect while communicating with me	Intonation of communication affects my performance	I would evaluate communication within the organisation as	I would evaluate communication within my department as	The interlocutor's approach to communication affects the understanding of the message
N	Valid	200	200	200	200	200
	Missing	0	0	0	0	0
Mean		2,66	3,77	2,77	3,08	4,18
Mode		2	4	2	3	4
Std. Deviation		1,254	1,011	1,244	1,022	,746
Minimum		1	1	1	1	2
Maximum		5	5	5	5	5

Table 2 Average levels of respondents' agreement with claims about respect in communication

		I feel understood and respected in communication within the organisation	Informal communication prevails within my department	I believe that professional communication is prevalent within the organisation	Colleagues treat me with respect when communicating with me	Communication within my department makes me feel heard
N	Valid	200	200	200	200	200
	Missing	0	0	0	0	0
Mean		2,77	3,26	2,86	3,18	3,13
Mode		2	3	3	3	3
Std. Deviation		1,284	1,057	1,280	1,154	1,125
Minimum		1	1	1	1	1
Maximum		5	5	5	5	5

In communication, regardless of its level or direction, it is crucial that all parties are respected and that actual dialogue, not monologue or arguing, is possible. The state of communication as a developed tool and part of business

processes is reflected in the results of Tab. 2. It is evident that respondents moderately agree with the claims on these issues. Communication within departments is such that respondents feel heard from their interlocutors (colleagues in the department and direct superiors), with an average agreement rating of 3.13. This rating is slightly higher for communication with colleagues in general (3.18), while at the organisational level overall - regardless of type or level of communication - the rating is 2.77. Within departments, informal communication between colleagues predominates, indicating a relatively relaxed atmosphere and contributing to a sense of connection among colleagues, but only if it also reflects mutual respect.

Table 3 Average levels of respondents' agreement with claims about the characteristics of communication and its impact on performance

		I feel safe to express my views and opinions	I often get interrupted and not heard until the end of my thought	Interrupting and ignoring my words affects my self-confidence and performance at work	I believe that everyone is heard and respected within the organisation regardless of their position	I have a feeling that some employees' opinions are belittled and ridiculed
N	Valid	200	200	200	200	200
	Missing	0	0	0	0	0
Mean		2,88	2,82	3,13	2,15	3,58
Mode		3	4	4	3	3
Std. Deviation		1,207	1,212	1,277	1,525	1,561
Minimum		1	1	1	1	1
Maximum		5	5	5	5	5

The level of agreement with the statement that all individuals in the organisation are heard and respected is relatively low (2.15). The situation is somewhat better when it comes to agreement with the statement that they are interrupted during communication and not listened to fully (2.82), as well as with whether they feel comfortable expressing their honest views and opinions (2.88). It can be concluded that employees feel somewhat freer and more relaxed when expressing their opinions, even though they are often not fully heard or acknowledged. This is a positive sign, as the presence of debate fosters progress, movement, and the discovery of solutions that benefit the entire organisation.

A reduced influence of arbitrariness positively impacts the quality of management and healthy interpersonal relationships. It is essential to highlight that the observed higher level of agreement with the statement that interruptions and the dismissal of ideas in communication affect employees' confidence and work efficiency (3.13) reflects the previously mentioned risk of conflict development and the diminished effectiveness of communication as a tool for managing processes and interpersonal relationships. It can be concluded that communication affects many areas - from knowledge acquisition and transfer to task execution and employee development, as well as tools for managing interpersonal relationships and business processes. Communication is a starting point and one of the key inputs for executing every organisational process. Therefore, it must be taken seriously and regulated not only through rules but also through established standards.

Table 4 Average agreement levels of respondents regarding the statements about the effect of various communication aspects and values on interpersonal relations in the organisation and its productivity

		Communication has a significant effect on knowledge transfer within the organisation.	I feel uncomfortable to ask questions when I don't understand something immediately.	When the work atmosphere is cosy, I am more efficient in my work	Gossiping within the organisation negatively affects interpersonal relations	To feel happy and fulfilled with my job, I need to have a comfortable work surroundings, good interpersonal relations and communication
N	Valid	200	200	200	200	200
	Missing	0	0	0	0	0
Mean		4,25	3,10	4,31	4,18	4,71
Mode		4	3	5	5	5
Std. Deviation		,854	1,319	,731	,890	,572
Minimum		1	1	2	2	2
Maximum		5	5	5	5	5

The importance of communication is evident from Tab. 4, as the levels of agreement with the statements indicate that it affects knowledge transfer within the organisation, the effectiveness and quality of interpersonal relationships, and ultimately, employee satisfaction since communication is part of the company atmosphere, organisational climate, and employee environment. Without good communication, knowledge within the organisation is not transferred correctly. In the long term, this leads to reduced employee effectiveness and increased staff turnover, as employees - especially younger ones - often leave more quickly from organisations where they cannot acquire quality knowledge, which reduces competitiveness and hinders the attraction of new skilled personnel.

Communication that hinders knowledge transfer is often practised by individuals, whether in leadership positions or not, who do not fully understand business tasks and processes or lack the skills to convey and communicate knowledge and practices effectively. As a result, during the training of new colleagues, they create confusion and shape new employees in a way that prevents their proper development. Such communication can make some individuals uncomfortable

about asking follow-up questions and, when dealing with egotistical interlocutors, even provoke resistance to questions being asked.

Employees on the same hierarchical level tend to understand each other better due to their position within the organisation, but they often highlight shortcomings in communication with their superiors. This is mainly so because superiors frequently use more complex language, adopt a more aggressive communication style, and fail to understand their subordinates, viewing them as counterparts in a "conflict" dynamic, where tasks are assigned, and specific results are expected. The extent to which this is present in the selected organisation is evident from the results in Tab. 5, where the level of agreement with the statement that superiors communicate unclearly is 3.75. There is a prevailing impression that such communication aims to obscure the understanding of the message (3.69).

Table 5 Average levels of agreement with claims about communication by superiors

		Superiors' communication is not clear enough.	The goal of superiors' communications seems not to be conveying the message.	Superiors attempt to be indirect and unclear in their communication if they themselves do not understand the problem.	I am not embarrassed to say that I do not understand and/or know something	Superiors sometimes do not choose their words in communication with subordinates
N	Valid	200	200	200	200	200
	Missing	0	0	0	0	0
Mean		3,75	3,69	3,22	3,29	3,42
Mode		5	5	4	3	4
Std. Deviation		1,460	1,609	1,488	1,369	1,342
Minimum		1	1	1	1	1
Maximum		5	5	5	5	5

Table 6 Average levels of agreement with claims about experiences in organisational communication

		Directly offensive communication is always verbal.	I have never been apologised to in cases of inappropriate communication	Superiors express their stress and nervousness in their communication manner	Communication towards HRM is not regulated to achieve the objective of this department	During my work life in this organisation, I experienced discrimination and humiliation
N	Valid	179	171	200	200	200
	Missing	21	29	0	0	0
Mean		3,71	4,29	3,79	3,06	3,35
Mode		5	5	5	4	4
Std. Deviation		1,504	1,291	1,566	1,179	1,476
Minimum		1	1	1	1	1
Maximum		5	5	5	5	5

Table 7 Median values of agreement with the first hypothesis' claims

		Communication significantly affects the transfer of knowledge within the organisation.	When the work atmosphere is pleasant, I am more productive.	To be satisfied and fulfilled with my job, I must have a pleasant working environment, good interpersonal relations and communication.
N	Valid	200	200	200
	Missing	0	0	0
Mean		4,25	4,31	4,71
Mode		4	5	5
Std. Deviation		,854	,731	,572
Minimum		1	2	2
Maximum		5	5	5

This gives employees the impression that even their superiors do not always fully understand the task or subject matter being discussed, making them less willing to contribute actively to task completion or problem-solving (3.22). It is also not uncommon for superiors to use inappropriate informal communication (3.42), including occasional swearing and seemingly polite language that shifts responsibility or belittles and discriminates by implying that someone has misunderstood or failed to grasp something correctly.

It should be pointed out that, more often than not, no one stops to question whether they have conveyed the message effectively or adequately addressed the issue with their interlocutor - precisely because communication is dominated by vanity and ego. A grounded and approachable communication style significantly impacts results and the development of organisational commitment and collegiality, and it encourages initiative and participation by communicating solutions to business-related tasks.

The organisation of communication is also a key factor in its effectiveness. Communication is generally structured through formal and informal channels, as well as digital and conventional means. In the selected organisation, communication occurs through both formal and informal modes. Formal communication is written (primarily submitted via e-mail and other digital communication tools such as Teams, SharePoint, DevOps, and similar digital platforms, etc.).

Informal communication is verbal or conducted via telephone/mobile phone and is characterised by conversational language. Such communication should not be used for inappropriate behaviour or targeted actions toward

employees. However, this is not always the case, as respondents agree with the statement that even slightly offensive communication tends to occur through informal communication and informal communication channels (3.71) and that they have experienced discrimination or humiliation in communication during their work career (3.35) (Tab. 6). This topic is sensitive and influenced by subjective impressions, so caution is needed when interpreting the results. Nevertheless, it is essential to note that in large organisations, it is challenging to control and monitor such aspects of communication and to manage so many employees with diverse personalities, experiences, and life circumstances. Some individuals transfer their private stress into the workplace, and, in case of dissatisfaction, they communicate with colleagues in a brusque and impatient manner.

Given that this is a large and highly successful organisation, it is a fact that stressful situations and pressure are sometimes prominent. In their pursuit of high goals, superiors may transfer this stress to employees, creating a pressuring approach to communication. This can occur both unconsciously and consciously. Regardless, and acknowledging the fact that the human resources department should monitor certain aspects, including communication, as a part of interpersonal relations, it is clear that managers and team leaders need to address this issue and use communication as a tool for achieving goals, not straying from them. This can be accomplished through management training and the development of communication skills.

4 DISCUSSION

Based on the research results, it has been observed that communication within the selected organisation shows certain shortcomings both in top-down and bottom-up directions. These issues stem partly from the organisation's strategic and operational practices, as lower-level employees lack participation in solving problems related to certain levels and tasks. The organisation's communication style is characterised by the use of informal channels, particularly in situations involving insecure interlocutors, as well as by the use of communication as a tool by superiors to assert authority over employees, which is an issue that is more pronounced in the retail sector. The ego dominates communication, and there is a lack of constructive discussion, which is an oversight on the part of human resources management, which does not place enough emphasis on communication as a tool for managing interpersonal relationships. Communication effectiveness is also reduced due to violating other rules and ethical principles (such as during hiring, favouritism, or protection of specific individuals). This highlights the need to define certain aspects of communication through explicit rules and standards and to embed the organisation's core values into its communication practices.

The results presented in Tab. 7 support the stated hypothesis, as there is an overwhelming level of agreement with the statements regarding the impact of communication on knowledge transfer, motivation, and efficiency. These interconnected categories produce synergistic effects, positioning communication as the primary tool for managing human resources, interpersonal relationships, and overall business operations.

Based on the analysis of the research results, employee interviews, and the analysis of internal communication structure, several weaknesses can be identified: lack of communication standards and norms; insufficient participation of employees (subordinates) in communication; limited use of communication channels; low investment in communication improvement; communication noise caused by ego; and minimal involvement of human resource management in overseeing interpersonal relationships through communication and its development.

The observed shortcomings and anomalies in communication within the researched organisations indicate a lack of adaptation to change and rapid growth, the increasing complexity of business processes resulting from a growing number of employees, and the accelerating changes and evolving needs of communication stakeholders.

The organisation's internal communication can be rated with a grade of 3. In eliminating its weaknesses, it is critical to capitalise on its communication strengths, which can be identified as follows: the predominance of informal communication, which can be utilised to create a pleasant atmosphere and help mitigate stress and pressure; mutual familiarity among employees and the ability of human resource management to effectively use communication tools to manage interpersonal relationships; an extensive network and easy access to communication equipment (mobile phones, laptops, telephones); and strong IT support in maintaining the quality of information and databases essential for communication and communication processes.

To improve communication, it is essential to seize specific opportunities to foster new strengths and sources of competitive advantage. Key among these are adopting new knowledge and practices in communication; educating employees on communication methods and the development of communication skills; using communication as a comprehensive set of tools for managing interpersonal relationships; training HRM personnel to manage interpersonal relations through communication and enhancing the quality of communication by aligning it with the organisation's core values.

In conclusion, it is essential and critical to eliminate weaknesses, as they undermine all efforts related to improvement and the realisation of opportunities. Weaknesses act like erosion, penetrating the communication structure; therefore, for a communication system and model to be sustainable, it must foster and reflect the organisation's core values and strategic goals.

5 CONCLUSIONS

Communication is a part of everyday life for both an individual and organisations. It is a broad concept encompassing practically everything that sends a specific message to the public, whether external or internal. Communication falls within the domain of the marketing

function and is the responsibility of human resource management and corporate communications. Corporate communications are responsible for the organisation's general external communication, while human resource management oversees employee internal communication. The communication style in the selected organisations is partly defined by function, as there is an established practice of a stricter and more aggressive top-down style, especially when directors are communicating. This style often produces the opposite effect on employees, as it can lead to feelings of oppression and workplace bullying. As such, this area is very sensitive and can become a source of misunderstandings, negatively affecting the development of loyalty and commitment to the organisation and potentially leading to the formation of cliques and the escalation of conflicts. The latter has a negative impact on healthy competition and motivation to perform, especially if it becomes widespread. This is typically the case with consistently strict and aggressive communication styles. However, when such a style is used selectively and combined with other approaches, it helps to balance relationships and supports the ability to manage interpersonal dynamics effectively.

Communication should serve as the input and output (feedback) of every business process and be part of the company's information system database, used and updated by all departments and units. In this way, human resource management can use such communication to effectively monitor, evaluate, and manage interpersonal relationships to achieve business goals. The integration of all strategies enables the development of a sustainable system. Interpersonal relationships are a dynamic variable, constantly subject to change, and best managed precisely through communication.

Communication is characterised by many elements, highlighting its complexity and the various possibilities it offers for managing interpersonal relationships. The organisation's values must be communicated to connect the organisation and the individual, align their efforts, and unify individual and organisational goals.

Although this research provides useful insights into the relationship between communication and organizational performance, some directions for future research can still be indicated, and they are based on the following: the research should be expanded to a larger number of organizations throughout Croatia, which would allow for variation in responses and thus reduce the limitations in generalizing the interpretation of the results. Furthermore, a longer time series of research is recommended in order to reduce the short-term nature of trends, i.e., future research should be expanded to a full-year timeframe. All of these recommendations aim to improve the credibility and application of research results on the impact of communication on organizational performance. The conclusions of such research would further highlight the potential for managing business operations through communication itself and, ultimately, through interpersonal relationships. These variables can significantly influence employee performance, as communication and interpersonal relations convey and transmit a range of intrinsic factors that motivate individuals within the organisation.

6 REFERENCES

[1] Žepić, M. (2000). *Latinsko-hrvatski rječnik*. Zagreb. Školska knjiga. (in Croatian)

[2] Lamza-Maronić, M., & Glavaš, J. (2008). *Poslovno komuniciranje.* Osijek. Sveučilište J. J. Strossmayera u Osijek. (in Croatian)

[3] Argenti, P. A. (2003). *Corporate Communication*. McGraw-Hill International Edition. 6th edition. New York.

[4] Kuran, M. L., & Jelić, P. (2014). *Poslovno komuniciranje*. Veleučilište "Marko Marulić" u Kninu. (in Croatian)

[5] Men, L. R., & Yue, C. A. (2019). Creating a positive emotional culture: Effects of strategic internal communication and its impact on employee supportive behaviors. *Public Relations Review, 45*(3), 101764. https://doi.org/10.1016/j.pubrev.2019.03.001

[6] Yue, C. A., Men, L. R., & Ferguson, M. A. (2021). Examining the effects of internal communication and emotional culture on employees' organizational identification. International Journal of Business Communication, 58(2), 169–195. https://doi.org/10.1177/2329488420914066

[7] Qin, Y., & Men, L. R. (2023). Exploring the impact of internal communication on employee psychological well-being during the COVID-19 pandemic: The mediating role of employee organizational trust. *International Journal of Business Communication*. Advance online publication. https://doi.org/10.1177/23294884231168794

[8] Welch, M., & Jackson, P. R. (2007). Rethinking internal communication: A stakeholder approach. Corporate Communications: An International Journal, 12(2), 177–198. https://doi.org/10.1108/13563280710744847

[9] Roberson, Q. M. (2021). Diversity in the workplace: A review, synthesis, and future research agenda. Routledge. https://doi.org/10.4324/9781003140518

[10] Buble, M. (2011). *Poslovno vođenje*. Zagreb. M.E.P. (in Croatian)

[11] Argenti, P. A., & Courtney, M. B. (2009). *Digital Strategies for Powerful Corporate Communications*. McGraw-Hill Education (India) Pvt Limited.

[12] Jacobs A. Z., & Watts D. J. (2021). A large-scale comparative study of informal social networks in firms. *Management Science, 67*(9), 5489-5509. https://doi.org/10.1287/ mnsc.2021.3997

[13] Jöhnk, J., Weißert, M., & Wyrtki, K, (2021). Ready or Not, AI Comes— An Interview Study of Organizational AI Readiness Factors, *Business & Information Systems Engineering*. 63(1), 5–20. https://doi.org/10.1007/s12599-020-00676-7

[14] Linjuan, R., Marlene, N., & Cen, A., (2020). Examining the Effects of Symmetrical Internal Communication and Employee Engagement on Organizational Change Outcomes, *International Journal of Business Communication*, 1–27. https://doi.org/10.1177/2329488420914066

[15] Zuzul, T., Pahnke, E., C., Larson, J., Bourke, P., Caurvina, N., Parikh Shah, N., Amini, F., Weston, J., Park, Y., Vogelstein, J., White, C., & Priebe, C., E. (2024). Dynamic Silos: Increased Modularity and Decreased Stability in Intraorganizational Communication Networks During the COVID-19 Pandemic. *Management Science, 71*(4), 3428-3448. https://doi.org/10.1287/mnsc.2022.02797

[16] Susita, D., Ketut Sudiardita R., Purwana, D., Wiradendi Wolor, C., & Merdyantie, R., (2020). Does organisational commitment mediate the impact of organisational culture and interpersonal communication on organisational citizenship behavior? *Management Science Letters, 10*, 2455–2462. https://doi.org/10.5267/j.msl.2020.4.004

[17] Marković, E., (2019). *Trendovi poslovne komunikacije u budućnosti* Sveučilište Josipa Jurja Strossmayera u Osijeku. Ekonomski fakultet u Osijeku, Osijek. (in Croatian)

Authors' contacts:

Andreja Rudančić, PhD, Associate Professor
(Corresponding author)
Libertas International University,
Trg J. F. Kennedy 6b, 10 000 Zagreb, Croatia
e-mail: arudancic@libertas.hr

Petar Mišević, PhD, Associate Professor
University North, University Center Varaždin,
104. brigade 3, 42000 Varaždin, Croatia
e-mail: pmisevic@unin.hr

Sandra Šokčević, PhD, Full Professor
Libertas International University
Trg J. F. Kennedy 6b, 10 000 Zagreb, Croatia
e-mail: ssokcevic@libertas.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20251022075850

Original scientific paper
Received: 2025-10-22, Accepted: 2026-01-08

Technological Innovation and Circular Economy: Smart Containers and Real-Time Monitoring for Efficient Waste Management

Milner Liendo*, Manuel Hilario, Alonso Tenorio, Jaime Cancho

Abstract: Inefficient municipal solid waste management represents a significant obstacle on the transition towards a circular economy in densely populated urban environments, such as the district of San Juan de Lurigancho, Lima, Peru. This study presents the implementation of advanced technologies, specifically smart containers integrated with sensors and real-time monitoring systems, to optimize waste collection. This resulted in decrements in operating costs and unrecycled waste by 20% and 27%, respectively, whilst increasing citizen engagement by 18% regarding segregation and recycling activities, highlighting positive changes in community behavior. However, the present technological infrastructure, technical training and necessary funds are all insufficient, which presents limitations on the deployment and scalability of these measures. This study thusly indicates the necessity to elaborate public policies, environmental education and engagement strategies to guarantee system stability. These findings both provide a replicable model to address these issues and offer a roadmap for the integration of smart technologies in sustainable city planning, bridging technological innovation and sustainability.

Keywords: circular economy; IoT sensors; waste management

1 INTRODUCTION

One of the greatest challenges in modern cities is proper waste management, only exacerbated by a quick population growth and following overstress of resources. It is estimated in the Global Waste Management Outlook 2024 [1] that municipal solid waste generation will increase from 2.3 billion tons in 2023 to 3.8 billion tons by 2050, being a 65% increase in less than three decades. This necessitates for collection, treatment and disposal infrastructures, especially in densely populated areas. In addition, improper management contributes to 5% of global greenhouse gas emissions, worsening the climate crisis [2]. This makes the implementation of smart technologies an unavoidable solution to mitigate these adverse effects. The integration of advanced sensors, real-time monitoring and route optimization algorithms has proven to better operational efficiency and reduce costs, minimizing environmental footprint. However, structural, financial and regulatory barriers persist, hindering the scalability of these innovations. This study analyzes the impact of these technologies on the sustainability of urban waste management and their role in the transition to an efficient and resilient circular economy.

These obstacles are also present in San Juan de Lurigancho (SJL), the most populous district in Lima with over 1.1 million inhabitants. The district faces a critical waste crisis, where 40% of residents perceive the current management as 'deficient' or 'very deficient'. Despite some environmental awareness, only 25% of the population consistently practices waste segregation at home. Furthermore, the existing infrastructure struggles with traditional collection routes that do not meet the demand of its high density. Addressing these challenges requires a shift from linear models to a circular economy, supported by the technological interventions proposed in this study to improve collection efficiency and community participation.

The circular economy changes the perception of waste management, promoting the reduction, recycling and reuse of materials. According to [3], it seeks to keep the value of materials and products for as long as possible, allowing production systems to reintroduce them to their cycles, generating less waste. To be effective, it should be complemented with smart technologies to optimize waste management, finally encouraging sustainable and efficient practices in densely populated urban areas, such as SJL.

Technological solutions, like smart containers and waste management systems with real-time monitoring, In this study, the system was implemented in San Juan de Lurigancho from March 2024 to February 2025, involving the deployment of 50 smart containers with a capacity of 1,100 liters each. To ensure precise monitoring, these units are equipped with HC-SR04 ultrasonic sensors controlled by an ESP32 microcontroller. The sensors emit high-frequency sound waves to measure the fill level in real-time. This data is transmitted via a GSM/GPRS wireless module using the MQTT protocol to a centralized cloud platform.



Figure 1 Data communication flow between smart containers and the management system

This architecture allows the system to visualize capacity and generate optimized collection routes based on actual container occupancy, replacing fixed schedules. These technological solutions prove to be effective in various studies. For example, [4] concluded that the use of sensors in containers favors the planning of collection routes and lowers operation costs. Likewise, [5] exposed that these technologies optimize efficiency and improve transparency in waste management, fostering trust and citizen participation.

Recent research [6] found that using Internet of Things (IoT), advanced algorithms and cloud analytics can bring significant improvements. Among these, there was a 32% more efficient routing, 29% less fuel consumption and 33% better performance in waste processing. The wide spectrum of effects these enhancements brought hint at the potential of these technologies to improve the economic, operational and

environmental areas of municipal waste management.

A circular economy could also help to build a sustainable and efficient waste management. However, in the context of urban settings like SJL, there are noticeable limitations to their adoption. These include economic and infrastructural constraints, along with limited technical expertise among operators. This commonly happens in developing countries, where the socioeconomic setting can restrict technological progress, requiring for thorough strategies that promote the integration, innovation and strengthening of technologies and institutions [7].

These results show the need to integrate advanced technologies in solid waste management to enhance sustainability and efficiency. However, the constraints found in places like San Juan de Lurigancho restrict proper implementation. These include reduced budgets, deficient infrastructure and poor operator training, remaining as obstacles to development. To address them, it's needed to procure innovative technologies, stronger institutions and increased community participation.

This article builds upon previous work about smart waste management by evaluating the implementation of modern technologies in densely populated urban areas with infrastructural and economical restraints, as is the case of San Juan de Lurigancho, a district of Lima in Peru. While most other studies focus places with better infrastructure or better funding, this research produces evidence on the integration of technological innovations, along with the participation of local communities, carefully designed education and the addition of supporting policies. With this, the study tends to gaps identified in other works and suggests a scalable approach that can be adapted to other urban settings under similar constraints.

The study's main goal is to assess the situation and suggest technological innovations for better management of solid waste in San Juan de Lurigancho, easing the transition to a circular economy. It focuses on sustainability and modern technologies, along with technical and economic feasibility for the local setting. It also designs strategies to connect innovation, technology, community engagement and institutional strengthening, with their union as a vital component of the circular model's success.

Beyond its practical application, its focus is building a scalable, replicable model that adjusts to other urban settings. With this, it warrants valuable improvements in environmental sustainability, operational efficiency and socioeconomic development, placing the foundations to implement public policies emphasizing an integrated waste management.

This study has a mixed design, with elements both quantitative and qualitative used to understand the environmental, economic and social implications of implementing innovative technologies for solid waste management. This hybrid design produces useful information for better decision making, supporting the optimization of urban waste collection and treatment.

On this research, the quantitative indicators selected for analysis include the reduction of unrecycled waste, advanced algorithms for route optimization and greater involvement of the community in separating and recycling waste. On the qualitative side, it gathers the perceptions and behaviors of local participants to discover opportunities and obstacles to implement these technologies.

The purpose of this design is to evaluate the possible outcomes and explore the interaction between emerging technologies and community dynamics. Municipal solid waste management should be thoroughly addressed by technological tools, environmental education and strategies for social participation [8].

The results of this study will help to design optimized pathways and strategies for more efficient operation for more sustainable and affordable practices. Its final goal is to create a management model that is replicable in other urban contexts, specifically by testing whether IoT-enabled bins can reduce the number of collection trips by 20% within six months and if awareness campaigns can increase the recycling rate by 10 percentage points within a year. These expected outcomes provide a clearer sense of the study's analytical focus, ultimately helping to transition into a circular economy.

Using a qualitative analysis will allow to collect subjective information, like the perspectives of the citizens. This allows us to detect possible practical obstacles and yields detailed information about waste management. With this, the study seeks to improve quality of life, lower negative environmental impacts and increase economic sustainability for San Juan de Lurigancho. This also supports one of the UN's Sustainable Development Goals [9], specifically procuring a sustainable consumption and production, through the design of compliant policies.

Using multi-objective methods to locate waste accumulation can significantly improve the performance of waste collection in actual urban conditions [10]. Similarly, optimizing micro-route sequencing and examining transfer station adequacy prove to be effective in Argentinian cities [11]. Because of this, improving the efficiency of waste management requires strategies that incentivize community engagement and institutional strengthening.

Circular economy policies are essential for optimum resource usage and the mitigation of pollution in developing cities [12]. In fact, this study not only looks for short-term improvements for the San Juan de Lurigancho population, but also sets the groundwork for public policies that can be replicated both nation-wide and internationally. Implementing sustainable waste management strategies has environmental and economic benefits, offering a scalable management model for regions under similar conditions.

2 LITERATURE REVIEW

In this research, there are included theories and procedures that support the connection between technological innovation, the circular economy and an efficient management of municipal solid waste. These are theoretical bases that support digitization, process optimization and adopting emergent technologies as crucial to develop sustainable waste management models. Some relevant factors for understanding include the following.

2.1 Circular Economy

In contrast to the usual linear manufacturing model, the circular economy allows a deep understanding of economic, social and economic sustainability. This model promotes the sustainable production and delivery of goods and services, lessening the negative effects of more traditional management systems while using resources more efficiently throughout their lifecycles [3]. Similarly, it is a structural change in how society views the environment by reducing resource scarcity and enabling sustainable growth via the optimization in the use of materials and their regeneration [13]. Implementing circular approaches is motivated by a persistent lack of raw materials for industry, caused by the overexploitation of resources and poor management.

The implementation of circular economies has effects on different scales. At the smallest level, they encourage sustainable business and consumer practices; at the medium level, they help in coordinating actors for a greater resource efficiency; and at the largest level, they influence national and local policies, enabling closed resource loops that minimize waste and increase industrial performance [14]. As a result, it's an essential component to address the current ecological crisis on a global scale and to develop sustainable and efficient industrial methods.

A central economy has three core principles that seek to transform the global production and consumption models into a regenerative and sustainable system. The first is the preservation and increase of natural capital with the efficient management of renewable resources and supporting technologies with a lower impact on the environment, thus aiding in the recovery of ecosystems and securing essential raw materials on the long term [13]. The second principle is resource optimization through product and process redesigns that comply with circularity criteria.

Recirculating material within biological and technical loops, through reuse and recycling, is vital to reduce waste and lengthen product lifetimes [14]. This reduces demand for virgin inputs and lowers the energy and environmental burdens of extraction and processing. The third principle emphasizes system-wide efficiency by limiting the social and environmental harms of economic activity, from reducing toxic substances and emissions to improving mobility, food, and health systems [13]. Applying these principles both quickens the transition to a more sustainable economy also allows innovation, investment and technological development that bolster urban resilience to emerging environmental and socioeconomic risks.

The circular economy has broad benefits across economic, environmental, business, and social dimensions. As noted in [3], it spurs growth by optimizing resource use, lowering production costs, and innovating areas like reverse logistics and energy efficiency. It also creates jobs in small and medium industrial enterprises, reinforcing resilience and competitiveness. Environmentally, circular strategies can cut CO_2 emissions by up to 50% by 2030, decrease raw-material extraction, and help mitigate climate change [3].

For firms, the model reduces supply chain upsets, lowers costs by recycling input, and deepens customer loyalty through engagement and responsible consumption [14]. For citizens, this model improves quality of life by extending product durability and discouraging planned obsolescence [13]. Beyond these general benefits, empirical studies in Latin America, such as those conducted in Argentinian and Peruvian urban centers, demonstrate that successful implementation depends on integrating local data with specific micro-route sequencing and community participation methodologies [11].

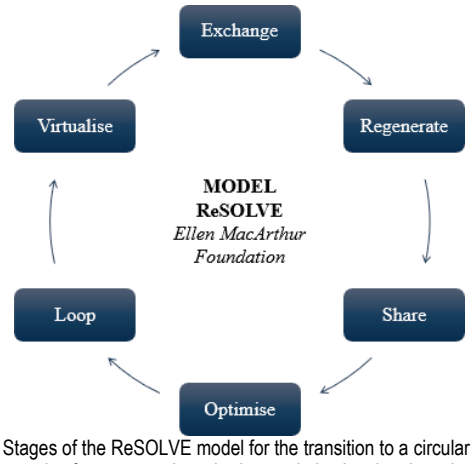


Figure 2 Stages of the ReSOLVE model for the transition to a circular economy, including strategies for regeneration, sharing, optimization, looping, virtualizing, and exchanging. Source: Ellen MacArthur Foundation [15]. Author's elaboration.

The ReSOLVE model, as designed by the Ellen MacArthur Foundation [3], has six core strategies to implement a circular economy, regenerating resources, optimizing processes and achieving digital transformation. The first one, regeneration, increases the use of renewable energies and the restoration of ecosystems by reintroducing biodegradable matter into natural cycles, increasing environmental resilience and limiting dependence on finite resources [14]. At the same time, sharing and optimizing strategies could promote a better management of assets and materials by placing reusability and long product lifespans as priorities to reduce waste.

Remanufacture and recycling, core elements of closed loop strategies, are critical for a continued circulation of materials and lowering the demand of raw input [13]. In the same way, digitalization transformed the management of consumption and waste by virtualizing services and processes, improving operational efficiency [3]. Lastly, phasing obsolete materials for more advanced options boosts innovation and business sustainability. As Mauricio Canu (in [3]) emphasizes, successful application of ReSOLVE hinges on integrating corporate social responsibility, eco-design, and adaptation to innovative processes, yielding a more efficient, sustainable, and competitive production model.

2.2 Efficient Waste Management Technology Solutions

Advanced technologies reshaped waste management by optimizing collection, treatment, and final disposal. Innovations span from end-to-end digitization to automated systems that polish operations for greater efficiency and sustainability. For example, smart sensors enable real-time tracking of container fill levels, improving route planning and cutting operating costs and collection-related emissions. In parallel, applying artificial intelligence (AI) to sorting has

increased the accuracy of identifying recyclables, raising recovery rate and reducing landfill disposal [7].

Automation in treatment facilities through robots and high-precision sorting has boosted processing efficiency, reduced human intervention and errors. Beyond operational gains, these tools support the shift to a circular economy, where waste becomes a resource that reenters production cycles. Adopting such technologies is critical to meet current environmental challenges and advance sustainable management practices [7].

Integrating technological solutions has also modernized core practices by creating more efficient and sustainable approaches. In collection and transport, sensors installed in containers permit live monitoring of fill levels, which optimizes routes and lowers both operating costs and CO_2 emissions. The use of electric and automated vehicles further improves efficiency and lessens environmental impacts. In treatment, anaerobic digestion facilities convert organic waste into biogas, generating renewable energy and reducing the volume of organics sent for disposal [16].

Furthermore, thermal technologies, like pyrolysis and gasification, break down solid waste for energy production while notably reducing the amount to be finally discarded. Using AI and robotics aided systems on recycling operations helps to efficiently sort materials, increasing resource recovery. By using specific procedures to recover valuable metals and treat hazardous components appropriately, e-waste recycling also evolved, minimizing environmental risks. When combined, these technologies have the potential to improve management systems and promote ecologically conscious behavior [16].

Using waste management technologies enhances the social, economic and environmental areas. This saves expenses and operating times by increasing efficiency at every step of waste-related operations, from collection to final treatment [17]. For instance, using sensors to track container fill levels enables effective route planning and lowers CO_2 emissions from transport. This shows that these solutions promote sustainability by reducing negative environmental effects, lessening waste generation and easing the transition to a circular economy.

Since these systems can be implemented in a variety of conditions, such as urban and rural areas as well as other nations, regardless of development, another perceived advantage of the circular economy is its adaptability [18]. Therefore, aside of increasing resource efficiency by promoting reuse and recycling, it minimizes environmental risks through the efficient treatment and recycling of waste, lessening pollution in soil, air and water. This also benefits public health because proper management avoids needless exposure to hazardous waste and reduces the risks of inadequate practices. Overall, integrating emerging technologies improves operations and has a significant positive impact on society and the environment [18].

2.3 Technological Solutions in a Circular Economy

Integrating innovative technology to enable a thorough use of resources through reuse and recycling is crucial for a shift from linear production models to a circular economy. This implies creating recyclable goods and materials and implementing intelligent systems for resource optimization and waste management. Incorporating sensors and IoT into products and infrastructure allows to accurately track and monitor materials, aiding in process optimization and improving decision making in circular models.

Additionally, gathering and analyzing large datasets can yield useful knowledge to improve resource management and decision making in circular systems. Data on trends, chances to improve and greater resource efficiency can be obtained through a thorough study, leading in cost savings and waste reductions [19]. Beyond process enhancements, these technologies support the development of new business models that adhere to sustainability and environmental responsibility, both principles of the circular economy.

Adopting innovative technologies in many areas is essential to transition into a circular economy. On this regard, eco-design revolves around designing things with consideration for their whole lifecycle, from material selection to recycling and reuse, lowering their environmental impact from the start. A known example is 3D printing, which is a critical tool for environmentally friendly construction, creating structures with recycled materials and noticeably cutting down on waste. Overall, these technologies contribute to a more sustainable and efficient production process while also encouraging innovation in business models that follow circular economic principles, fostering a more responsible resource management and reducing their environmental impact [20].

2.4 Digital Transformation and Technology Solutions

Digital transformation is a crucial component of the modern business' competitive advantage, rearranging their operations and consumer relationships. Digital initiatives should equally prioritize social responsibility, environmental preservation and efficient economics in order to be truly sustainable [21]. Creating sustainable technology practices promotes balanced development, reduces pollutant emissions and helps to preserve natural resources. This means that business should consider the technology they choose and how it impacts society and the environment. They should ensure that, apart from growth or profits, their digital strategies take safety and wellbeing into account.

2.5 Waste Management and Public Policy

Managing municipal solid waste is a complex policy area that demands strong public involvement to encourage sustainable practices and improve waste management infrastructure. Integrated solid waste management is an essential part of reducing the contamination of water bodies, like Ecuador's Daule River, when supported by sensible regulations [22]. According to [23], management models that integrate waste reclamation can enhance environmental sustainability by encouraging reduction, reuse and recycling in accordance with the principles of a circular economy.

Additionally, there is a host of issues with urban waste systems, including biophysical, technological, social and administrative concerns that compromise public health and

the environment [24]. Overcoming them necessitates procuring a robust infrastructure and a greater public awareness about sustainable behaviors. Crucial actions include the formalization of informal sectors and the promotion of investments for material reuse. In Mexico, it was found that solely practical developments lose effectiveness, even with legal progress, showing gaps between policy design and execution, marking the need for stronger enforcement and monitoring [25].

Recent developments in solid waste management have been directed toward sustainable and efficient practices. Projections [26] show an increase of municipal solid waste from 2.3 billion tons in 2023 to 3.8 billion tons in 2050, urging immediate action. In this context, circular economy presents a viable route via product and material maintenance, and the regeneration of natural systems, thus advancing environmental and economic sustainability [19].

The design of policies should be followed by innovation in technology. Digitized processes and sensor-based data improved collection, transport, and treatment efficiency while reducing operating costs and pollutant emissions. Aligning public policy, technological adoption, and sustainable behaviors is essential to meet current Municipal solid waste (MSW) challenges and move toward a more sustainable future.

2.6 Corporate Sustainability

Sustainable development entails satisfying present needs without risking those of future generations by balancing social, environmental, and economic outcomes. This poses challenges for governments as they reconcile the use of natural and social resources with growth, and it hints at the interdependence of people, ecosystems, and economic systems. Growing a sustainability culture that fuses social, economic, and environmental values is therefore critical. Corporate sustainability implies actions that protect the environment and create stakeholder value without sacrificing profitability. It is important for governance mechanisms to embed sustainability into corporate strategy [27]. Today, executives increasingly recognize its relevance for brand image and competitiveness.

Consumers favor environmentally responsible firms, as seen in growing demand for eco-friendly products and green marketing. Thus, companies must adopt practices that balance profit with environmental protection and social responsibility to remain competitive and socially credible.

Business sustainability rests on three interlinked pillars. The environmental dimension reduces operational damages through actions such as cutting greenhouse-gas emissions, managing waste efficiently, and conserving natural resources [27]. The social dimension concerns the well-being of employees, customers, and communities through fair labor conditions, diversity and inclusion, and contributions to social development [28]. The economic dimension focuses on long-term profitability, shareholder value, and financial resilience [29].

3 METHODOLOGY

This research has a descriptive design, allowing to assess and comprehend waste management phenomena without altering the normal circumstances. This design works best to extract organized, precise information from existing variables, allowing to track behaviors and identify patterns. Descriptive work observes, documents and examines events on-site, in contrast to experimental research, which interacts with variables to evaluate changes.

In order to fully evaluate the current state of affairs, this study employs a mixed-methods approach. Data was gathered via a structured survey and semi-structured interviews. The survey instrument comprised 25 items divided into five sections: demographic data, waste management perception, use of technology, circular economy practices, and general opinions. To ensure reliability, the instrument was validated using Cronbach's Alpha (Tab. 5), yielding a high overall internal consistency of 0.84. Additionally, semi-structured interviews were conducted with key stakeholders to provide qualitative depth. The complete questionnaire, its statistical validation, and the summarized responses are provided in Appendix A. This enables the observation of trends and patterns in the use of technology for waste management in cities, producing proof to aid in understanding and creating sustainable strategies. Findings from descriptive studies serve as foundation for subsequent correlational or experimental research in virtue of offering a thorough overview of the phenomenon under investigation, yielding tools for well-informed decision making in the fields of circular economy and urban sustainability.

The goal of this fundamental or basic type of research is to expand the understanding of how innovative technologies improve management without actually putting them into practice. The scope is explanatory, identifying cause–effect relationships between smart waste-management systems and outcomes such as operational efficiency, reduced non-recycled waste, and citizen participation in separation.

The design is relational–causal: independent variables—including real-time monitoring and collection-system digitization—are linked to dependent variables such as route optimization and reduced environmental pollution. This framework generates empirical evidence on the feasibility of these systems in dense urban settings with limited waste-management infrastructure.

The target population comprises residents of San Juan de Lurigancho (Lima, Peru), with emphasis on areas experiencing the greatest management challenges. SJL is Peru's most populous district, with an estimated 1,245,145 inhabitants in 2024 [30]. High density intensifies management pressures as per-capita waste generation is substantial. These issues affect health, the environment, and quality of life, reinforcing the need to analyze local waste dynamics to propose effective, sustainable solutions.

Sample size for the finite population of 1,245,145 was calculated using a 95% confidence level and 5% margin of error, yielding 384 required respondents to ensure representativeness. The sample size (n) was determined using the formula for a finite population:

$$n = \frac{N \cdot Z^2 \cdot p \cdot q}{e^2 \cdot (N-1) + Z^2 \cdot p \cdot q}. \quad (1)$$

Where: n - Required sample size, N - Total population size (1,245,145), Z - Confidence level score (95%), e - Margin of error (0.05), p - Probability of success (0.5), q - Probability of failure (0.5).

Stratified random sampling was applied by dividing the district into homogeneous strata—e.g., residential, commercial, and infrastructure-limited areas—so that each subgroup is adequately represented. This strategy captures population heterogeneity and enhances the generalizability of findings across subgroups.

A mixed-methods design is employed. Structured surveys quantify waste-management practices among households and businesses, while semi-structured interviews explore operational challenges faced by authorities and service providers. Instruments include questionnaires assessing practices and attitudes toward circular economy concepts, followed by interview guides for municipal officials and collection companies. The qualitative tool probes policy implementation, operational obstacles and perceived effectiveness of current strategies. Combining quantitative and qualitative evidence yields a thorough, nuanced view of the issue and potential solutions.

Quantitative data from structured questionnaires were analyzed in IBM SPSS v.27 using descriptive statistics, frequency distributions, chi-square tests, and logistic regression to identify significant associations and predictive factors. Qualitative interviews were audio-recorded, transcribed verbatim, and coded thematically in ATLAS.ti v.9 to deepen interpretation of quantitative results.

4 RESULTS

Peru's municipal solid waste landscape has substantial barriers to sustainable development and circular-economy advancement. Regarding national statistics, the 'Informe Anual de Residuos Sólidos 2021–2022' [31] estimates a daily generation of approximately 21,000 tons of municipal waste nationwide, which averages to 0.8 kg per person daily. While other studies on urban waste models in IoT-enabled environments [32] suggest that per capita generation can vary significantly depending on the efficiency of collection systems and urban density, this study adopts the official national baseline of 0.8 kg/day to maintain consistency with the data provided by the Peruvian government [31].

This surpasses both managerial and treatment capabilities, particularly when considering current limitations on infrastructure. The report also shows that merely a 50% of waste is properly disposed of in sanitary landfills, with the rest being left in informal dumps, risking public health and the environment. These patterns raise the urge to reinforce policy and management initiatives that target reduction, reuse and recycling, following circular economy principles. Effective action in this regard is critical to reduce the harm of accumulation and improper disposal and to enable the sustainable development of the country.

Table 1 Solid Waste Generation in Peru

Year	Waste Generated (million tons)	Waste per Person (kg/day)
2018	4.8	1.4
2019	5.0	1.5
2020	5.2	1.5
2021	5.5	1.5
2022	5.7	1.6

Source: Prepared by the authors based on the 2024 field study.

As illustrated in Tab. 1, solid waste generation in Peru exhibited a continuous upward trend between 2018 and 2022. This growth is characterized by a steady increase in both total annual volume and the daily per capita output rate. The data reflects an intensifying pressure on urban sanitation systems, highlighting the urgent need for integrated management strategies to decouple population growth from waste production. This rise in waste production tonnage is indicative of both population expansion and increased economic and commercial activities.

This growth shows that the infrastructure for waste treatment and disposal not enough, urging for stricter policies that focus effective management, recycling and reduce non-biodegradable materials. It also requires advancing circular economy strategies that integrate greater participation from both public and private actors as well as technological innovation. On the long term, this would improve sustainability and lessen the environmental impact of Peru's waste system, urging to replace current structures with more effective and sustainable ones.

Deploying smart systems is thus an essential approach to increase operational efficiency and reduce environmental risks in modern cities. According to [4], containers with fill level sensors can reduce the cost of waste 20% by optimizing truck routes. Other recent studies show that implementing Big Data and Internet of Things (IoT) tools can improve real-time monitoring and streamline logistical operations, preventing excessive buildup of waste and lowering health hazards in dense urban areas [32].

Furthermore, incorporating artificial intelligence onto them allows to predict waste generation patterns, improves resource allocation and reinforces circular economy principles. Digitizing waste management procedures is a solution to the issues posed by the rise in waste generation in urban environments and is crucial for their sustainability.

Table 2 Impact of Smart Containers on Waste Collection

City	Cost Reduction (%)	Collection Trip Reduction (%)
Ciudad A	20	30
Ciudad B	15	25
Ciudad C	25	35

Source: Prepared by the authors based on the 2024 field study.

Collectively, Tab. 2 demonstrates that combining IoT sensors, AI, and real-time analytics optimizes routing and prevents needless trips. These technologies advance sustainability by reducing environmental impact and aligning with circular economy principles, urging the adoption of digitization to build more resilient urban models.

Promoting citizen participation is essential for effective municipal solid waste management. Initiatives that improve environmental education and awareness can transform behaviours and promote appropriate waste segregation. The results presented in Tab. 3 are based on the 2024 field study

conducted in San Juan de Lurigancho, which utilized a sample size of 384 participants to ensure a 95% confidence level. Data was collected through quantitative surveys to evaluate how different outreach strategies influence waste management perceptions. As noted in [33], organized educational programs can increase recycling rates significantly, suggesting a direct correlation between eco-friendly behaviour and structured awareness.

Furthermore, the data proves that keeping communities informed through mobile apps and digital platforms enhances collection efficiency. For consistent success, the collaboration between the public sector and civil society—supported by the financial incentives suggested by the trends in Tab. 3 is vital to making urban waste systems resilient to environmental changes.

Table 3 Effect of Awareness Campaigns on Recycling Rate

Type of Campaign	Increase in Recycling Rate (%)
Campaigns in Schools	50
Community Campaigns	40
Social Media	30
Local Media Advertising	35

Source: Prepared by the authors based on the 2024 field study.

The results presented in Tab. 3 are derived from the project's field study conducted in 2024. To clarify the methodology, the recycling rate and impact levels were estimated through a quantitative approach using a sample size of 384 participants from the San Juan de Lurigancho district, ensuring a 95% confidence level. The data collection involved questionnaires and registration forms to evaluate how different outreach strategies influence waste segregation habits. The findings highlight that educational programs in schools are the most significant drivers for behavioural change, aligning with the high awareness levels and the priority given to environmental education reported in the study's surveys. This underscores the importance of combining formal education with digital and community-based strategies to foster a long-term recycling culture.

As a strategy, the circular economy optimizes waste management and reduces the environmental impact of traditional linear production and consumption models. Implementing this model helps to lessen landfill waste by 30%, improving environmental sustainability and procuring efficient resource usage [34]. It considers material regeneration, maximal use of inputs and waste reduction through eco-design, reuse and advanced recycling.

Additionally, recent studies consider the adoption of a circular economy, beyond its positive impact on waste reduction, as capable of creating economic benefits via reducing production costs and promoting the design of new sustainable business models. Moreover, integrating modern technologies, such as data analysis in real time and artificial intelligence, allows for a more efficient and predictive waste stream management. This exposes the need to design public policies and regulatory frameworks that lead to the transition to this model, warranting a scalable, applicable model in various industrial and urban sectors.

The introduction of circular economy principles led to significant improvements in the reduction, reuse, and recovery of urban waste. As illustrated in Tab. 4, the transition from a linear to a circular model resulted in a substantial decrease in waste sent to landfills, which were previously characterized by high levels of inefficiency and contamination. This shift was accompanied by a notable increase in recycling rates, reflecting a strengthened capacity for material recovery. From an economic perspective, these strategies improved financial performance by increasing management cost savings, demonstrating that circular approaches are not only environmentally beneficial but also economically viable. These results position the circular economy as a critical factor for transforming waste systems, enhancing operational effectiveness, and promoting long-term sustainability in cities facing modern environmental challenges.

Table 4 Impact of the Circular Economy on Waste Management

Indicators	Before Implementation	After Implementation
Waste sent to landfill (%)	70	40
Materials recycled (%)	20	50
Management cost savings (%)	10	25

Source: Prepared by the authors based on the 2024 field study.

There are severe issues with waste management in Peru and the globe, requiring innovative, long-term solutions. Some critical measures to lessen the environmental effects of poor management include using modern technologies, progress in circular economy and increased public involvement. The findings prove that smart waste containers and continuous monitoring improve logistical planning, lower operation costs and enhance waste collection procedures. Additionally, circular economy practices have increased recycling and recovery rates while lowering landfill use. However, this success needs institutional backing as well as government programs that focus environmental education and investing on infrastructure. To achieve this, it is critical to foster innovation, improve regulation and increase public engagement in sustainable practices. Creating thorough and collaborative strategies is needed to implement a system that is efficient and flexible enough to respond to sustainable development goals.

5 DISCUSSIONS

The impact of the proposed study has been evident in that the integration of smart containers and the use of real-time tracking systems has greatly improved the management of waste disposal. The 20% reduction in the cost of operations experienced by the San Juan de Lurigancho district affirms the claims by Zaman (2014) that zero-waste management approaches strictly depend upon the use of real-time information to eliminate logistical waste. Contrary to the previous method, where operations were conducted using the static route method, the use of IoT sensors has ensured the use of dynamic collection, resulting in savings in fuel, in addition to being environmentally friendly by reducing the carbon footprint of the urban buses.

For engagement in citizenry, analysis of the data rendered in Tab. 3 indicates the highest efficiency in promoting change of behaviour through educational campaigns in schools. The observations from the table are in

line with the postulates put forth by Groot and Steg in 2009 on the effectiveness of institutional methods in promoting youth to adopt pro-environmental behaviours compared to more aged populations. Notwithstanding the 30% rise in the ambassador for increased knowledge regarding recycling initiatives for digital engagements (social and mobile initiatives), the gap between the two methods indicates the potential need for digital initiatives in boosting outcomes in regions with digital divides in cities like San Juan de Lurigancho.

Moreover, the financial feasibility of the circular economy paradigm, as denoted in the cost savings represented in Tab. 4, indicates that the changeover from the traditional linear 'take, make, dispose' to a circular economy paradigm represents a financial opportunity for developing urban centers. These observations signify that the government policy agenda should target technology investment subsidy schemes as well as reward mechanisms for the recycling processes. As pointed out in contemporary literature reviews, the convergence of the aforementioned systems calls for a collaborative mechanism where the government, technology companies, and citizens operate in the same cycle for ensuring the long-term sustainability of the processes via scalability mechanisms.

5.1 Research Limitations and Future Work

However, this research has some limitations that need to be discussed, as recommended by the reviewers. Firstly, although it is statistically representative with a sample size of 384 participants and a confidence level of 95% for the whole district, it may not take into account all of the socio-economic variables of the population. Moreover, since it is a survey, it might have a response bias since people have a tendency to over-report their pro-environmental behaviour for 'social desirability' reasons.

In fact, this experiment took place for a particular time during 2024, which means that there are factors involving seasonal changes in trash production, like plastic trash as well as organic trash production, which takes place around holidays. Longitudinal studies can help in determining how long such behaviour changes can occur for many years to come. Another aspect that can help in providing a stronger foundation for urban sustainability would involve exploring how new technology can help in trash tracking or trash production predictions.

6 CONCLUSIONS

Based on the above discussion and the data collected, this research concludes the following: The implementation of smart technologies, such as IoT sensors and real-time monitoring, serves as an integral triggering factor in the transition to a circular economy in densely populated areas like San Juan de Lurigancho. The first finding of this research reveals a 20% decrease in the operational cost of the municipality owing to the implementation of smart containers and effective routing algorithms, which proves the supremacy of data-led logistics over the fixed schedule method. The second finding reveals a 50% increase in the effectiveness level of educational campaigns conducted on schools, determining the age group of children to have the utmost impact in bringing a behavioural transformation among households. The third finding illustrates a 30% decrease in the dependency on landfills, thus effectively halting the danger of environmental pollution.

Secondly, with regards to social dynamics, it was found that a change in behavior is most effectively achieved through educational campaigns in schools, posting a 50% improvement over traditional media. This supports that young citizens play a role as "environmental ambassadors" among family members, initiating a bottom-up change in family practices regarding waste segregation. Finally, the systemic implementation of circular philosophy, using these technologies, was found to decrease landfill reliance by 30% and promote a more resilient city model with a focus on post-use material recovery rather than mere disposal.

On the strength of these findings, it would appear that cities should move from their traditional role of 'waste collectors' to 'resource recoverers'. For this, a two-pronged approach would be necessary: infrastructure development spending in digital form for facilitating a continuous flow of data, and integrating education on the ecosystem in the school education system of the civic body. Notwithstanding these findings, some natural constraints exist in the form of a sample size of 384 participants and the possibility of biased responses in the survey form of data, since the latter relies on attitudes and not actions.

Table 5 Validation of the Questionnaire by Cronbach's Alpha

Questionnaire Section	Item	Cronbach's Alpha value	Interpretation
Section 1: Waste Disposal Practices	Items 1-5	0.85	High reliability: Good internal consistency
Section 2: Knowledge about the Circular Economy	Items 6-10	0.8	Good reliability: Adequate internal consistency
Section 3: Recycling Practices and Responsible Disposal	Items 11-15	0.88	High reliability: Good level of consistency
Section 4: Attitudes towards Innovation in Waste Management	Items 16-20	0.83	High reliability: Satisfactory internal consistency
Section 5: Opinion on Waste Management and the Circular Economy	Items 21-25	0.79	Good reliability: Acceptable internal consistency
Questionnaire Total	All items (1-25)	0.84	High overall reliability

Source: Prepared by the authors based on the 2024 field study.

This study, therefore, concludes with the recommendation for the implementation of a large-scale pilot project in a controlled residential zone to test a "Pay-As-You-Throw" system. This model should be based on precision weight sensors and user-identification tags as main KPIs for rewarding waste reduction through tax incentives. Further, a 12-month longitudinal follow-up is recommended in order to monitor the durability of these behavioural changes. This model would be able to scale up effectively in other metropolitan areas in Lima and in similar emerging cities in the global south by integrating some emerging tools, such as Artificial Intelligence for predictive waste generation and Blockchain for transparent material tracking, hence assuring

long-term urban sustainability and environmental justice.

7 REFERENCES

[1] United Nations Environment Programme and International Solid Waste Association. (2024). *Beyond an age of waste: Global Waste Management Outlook 2024*. UNEP. https://www.unep.org/resources/report/global-waste-management-outlook-2024

[2] Naciones Unidas (ONU). (2023). Actúa ahora: Datos y cifras sobre la crisis climática. Recuperado el 30 de enero de 2024, de https://www.un.org/es/actnow/facts-and-figures (in Spanish)

[3] Lehmann, L. (2019). Economía circular: El cambio cultural: De la gestión de residuos a la gestión de recursos. Prosa y Poesía Amerian Editores. (in Spanish)

[4] Godoy, D. A., Sosa, E. O., Luft, R., Sosa, D., Belloni, E. A., & Benítez, J. D. D. (2015). Ciudades inteligentes: optimización en la recolección de contenedores de residuos domiciliarios. In XVII Workshop de Investigadores en Ciencias de la Computación (Salta, 2015). (in Spanish)

[5] Marín, E. N. (2024). Manejo de residuos sólidos en zonas urbanas en América Latina. *Visión de Futuro, 28*(2), 78-97. (in Spanish) https://doi.org/10.36995/j.visiondefuturo.2024.28.02.003.es

[6] Addas, A., Khan, M. N., & Naseer, F. (2024). Waste management 2.0 leveraging internet of things for an efficient and eco-friendly smart city solution. *PLoS ONE, 19*(7), e0307608. https://doi.org/10.1371/journal.pone.0307608

[7] Amores, M. S. S. (2023). La exigencia de una gestión eficiente de los residuos municipales en el contexto de la estrategia de economía circular. *Actualidad Jurídica Ambiental*, 1-33. (in Spanish) https://doi.org/10.56398/ajacieda.00278

[8] Gómez, J. B., & Bardales, J. M. D. (2020). Gestión de Residuos Sólidos Urbanos y su Impacto Medioambiental. *Ciencia Latina Revista Científica Multidisciplinar, 4*(2), 993-1008. (in Spanish) https://doi.org/10.37811/cl_rcm.v4i2.135

[9] Organización de las Naciones Unidas. (2015). Transformar nuestro mundo: La Agenda 2030 para el Desarrollo Sostenible. Naciones Unidas. Recuperado de https://unctad.org/system/files/official-document/ares70d1_es.pdf (in Spanish)

[10] Rossit, D. G., Toutouh, J., & Nesmachnow, S. (2020). Exact and heuristic approaches for multi-objective garbage accumulation points location in real scenarios. *Waste Management, 105*, 467-481. https://doi.org/10.1016/j.wasman.2020.02.016

[11] Molfese, S., Rossit, D., Frutos, M., & Cavallin, A. (2022). Optimization of waste collection through the sequencing of micro-routes and transfer station convenience analysis: an Argentinian case study. *arXiv*, 1(1), 1-34.

[12] Liu, Q., Xie, M., Mi, G., & Yuan, X. (2025). Integrating circular economy policies in urban policy making: Strategies for green energy and environmental management to reinforce city image in China. *Journal of Environmental Management, 375*, 124128. https://doi.org/10.1016/j.jenvman.2025.124128

[13] Belda, H., I. (2018). *Economía circular: Un nuevo modelo de producción y consumo sostenible*. Editorial Tébar Flores. (in Spanish)

[14] Arnedo, R., Jaca, C., León, C., & Omaetxebarria, M. (2020). *Guía práctica para implementar la economía circular en las pymes*. AENOR. (in Spanish)

[15] Almeida-Guzmán, M., & Díaz-Guevara, C. (2020). Economía circular, una estrategia para el desarrollo sostenible. Avances en Ecuador. *Estudios de la Gestión Revista Internacional de Administración, 8*(1), 35-57. (in Spanish)

[16] Araujo-Flores, A., & Delgado-Inga, O. (2021). Manejo sostenible de residuos sólidos domiciliarios: recolección y transporte al relleno sanitario. Caso cantón Naranjal. *Universidad-Verdad, 1*(78), 62–95. (in Spanish) https://doi.org/10.33324/uv.v1i1.364

[17] Amaro, R. E., Lengua, I. L., & Miguel, S. E. (2023). Sostenibilidad en la gestión de residuos de paneles solares en España. *Proceedings INNODOCT/19. International Conference on Innovation, Documentation and Education*, 413-420. (in Spanish) https://doi.org/10.4995/INN2022.2022.15786

[18] Meireles, R. R., Morales, G. P., Pereira, F. M., Nascimento, M. M. M., Duarte, J. P. M., Da Silva, D. F., & Leal, M. (2022). Avaliação de Impactos Ambientais dos Resíduos Sólidos na Universidade do Estado do Pará, Campus V - CCNT. *Revista Brasileira de Geografia Física, 15*(3), 1336–1351. (in Spanish) https://doi.org/10.26848/rbgf.v15.3.p1336-1351

[19] Navarro, A. C., Ferreira, D. H. L., Sugahara, C. R., & Conti, D. de M. (2021). *Economía circular: um estudo bibliométrico Revista Visão: Gestão Organizacional, 10*(2), 17–23. (in Spanish) https://doi.org/10.33362/visao.v10i2.2754

[20] Souza, J., & Silva, E. A. (2024). Um diálogo a partir das concepções e relações presentes entre economia solidária, empreendimentos solidários, inovações e tecnologias: *A dialogue based on the present conceptions and relationships between solidarity economy, solidarityenterprises, innovations and technologies*. *Revista Visão: Gestão Organizacional, 13*(2), e3330-e3330. https://doi.org/10.33362/visao.v13i2.3330

[21] Moreno-Felices, P., Puebla-Martínez, B., & Gelado-Marcos, R. (2021). Análisis, medición y control en el proceso de transformación digital de medios: Media Analysis Tool (MAT). *Tripodos, 46*, 157-178. (in Spanish) https://doi.org/10.51698/tripodos.2020.46p157-178

[22] Santillán-López, L. G., & Gines-Tafur, E. (2024). Gestión integral de los residuos sólidos y las políticas públicas en la mitigación de la contaminación del río Daule, provincia del Guayas, República del Ecuador. *Revista Científica Arbitrada De Investigación En Comunicación, Marketing Y Empresa REICOMUNICAR*, 7(14 Ed. esp.), 34-51. (in Spanish)

[23] Ortiz Castro, G., & Contreras Rivera, R. J. (2023). Políticas públicas, gestión de residuos sólidos municipales y sostenibilidad ambiental. Una revisión. *Ciencia Latina Revista Científica Multidisciplinar, 6*(6), 12258-12278. (in Spanish) https://doi.org/10.37811/cl_rcm.v6i6.4253

[24] Banco Interamericano de Desarrollo. (2023). Gestión de residuos urbanos para ciudades más limpias: Hacia una economía circular en LAC. Banco Interamericano de Desarrollo. Recuperado de https://blogs.iadb.org/ciudades-sostenibles/es/gestion-de-residuos-urbanos-para-ciudades-mas-limpias-hacia-una-economia-circular-en-lac/ (in Spanish)

[25] García-Mondragón, D., Cervantes-Zepeda, I., Gómez-Demetrio, W., Gallego-Alarcón, I., García-Pulido, D., & González-Blanco, G. (2023). Gestión de los residuos sólidos en México: análisis cualitativo de los diagnósticos básicos. *Inter disciplina, 11*(30), 215-242. (in Spanish) https://doi.org/10.22201/ceiich.24485705e.2023.30.81788

[26] Programa de las Naciones Unidas para el Medio Ambiente. (2024). El fin de la era de residuos: transformación de la basura en recursos. Recuperado de https://www.unep.org/es/noticias-y-reportajes/comunicado-de-prensa/el-mundo-debe-superar-la-era-de-los-desechos-y (in Spanish)

[27] Orquera, M. F. (2023). Aporte de las empresas ecuatorianas para alcanzar los Objetivos del Desarrollo Sostenible. *Universidad-Verdad, 2*(83), 66–87. (in Spanish) https://doi.org/10.33324/uv.v2i83.699

[28] Rosas, L. F., Ramírez, J. H., & Atlahua, A. D. (2024). Determinación del nivel de apego a la Responsabilidad Social Empresarial en microempresas de servicios en Escárcega,

Campeche, México. *European Scientific Journal, ESJ, 26*, 1. (in Spanish) https://doi.org/10.19044/esipreprint.2.2024.p1
[29] Cordeiro, S. L., Hulse, L., Martins, A. A. M., & Lehmkuhl, M. de S. (2021). O percurso histórico da sustentabilidade, suas dimensões e objetivos de desenvolvimento sustentável. *Professare, 10*(1), e2922-e2922. (in Spanish) https://doi.org/10.33362/professare.v10i1.2922
[30] INEI. (2024). Población de la provincia de Lima supera los 10 millones 292 mil habitantes. Recuperado de https://m.inei.gob.pe/prensa/noticias/poblacion-de-la-provincia-de-lima-supera-los-10-millones-292-mil-habitantes-14869/ (in Spanish)
[31] Ministerio de Vivienda, Construcción y Saneamiento del Perú. (2022). Informe anual de residuos sólidos 2021-2022. Gobierno del Perú. Recuperado de https://www.gob.pe/institucion/vivienda/informes-publicaciones/4045788-informe-anual-de-residuos-solidos-2021-2022. (in Spanish)
[32] Hussain, I., Elomri, A., Kerbache, L., & Omri, A. E. (2024). Smart city solutions: Comparative analysis of waste management models in IoT-enabled environments using multiagent simulation. *Sustainable Cities and Society, 103*, 105247. https://doi.org/10.1016/j.scs.2024.105247
[33] Carbonell-Alcocer, A., Romero-Luis, J., Gértrudix-Barrio, M., & Borges-Rey, E. (2022). Educating for a sustainable future through the Circular Economy: Citizen Involvement and Social Change. *Comunicar, 30*(73), 21-32. https://doi.org/10.3916/C73-2022-02
[34] Murazzo, L. F. (2023). Sistema de manejo de residuos sólidos y cuidado del medio ambiente de la empresa Minera Horizonte S.A. en el departamento de La Libertad - 2020 [Tesis de maestría, Universidad Nacional Federico Villarreal]. Recuperado de https://hdl.handle.net/20.500.13084/6884 (in Spanish)
[35] Melendez, J. R., Delgado, J. L., Chero, V., & Franco-Rodríguez, J. (2021). Economía Circular: Una Revisión desde los Modelos de Negocios y la Responsabilidad Social Empresarial. *Revista Venezolana de Gerencia, 26*(6 Edición Especial), 560-573. (in Spanish) https://doi.org/10.52080/rvgluz.26.e6.34

Authors' contacts:

Milner David Liendo Arévalo, Mg.
(Corresponding author)
Escuela de Posgrado Newman,
Av. Bolognesi 987, Tacna 23001, Peru
milnerdavid.liendo@epnewman.edu.pe

Alonso Tenorio Trigoso, Dr.
Universidad Nacional de Ingeniería
Av. Túpac Amaru 210, Rimac 15333, Perú
(511) 997059777, atenoriot@uni.edu.pe

Francisco Manuel Hilario Falcón, Dr
Universidad Nacional Federico Villarreal, Graduate School
Jr. Prolongations Camaná N° 1014, Lima 01, Perú
(511) 975871089, fhilario@unfv.edu.pe

Jaime Antonio Cancho Guisado, Mg.
Universidad Nacional Federico Villarreal, Graduate School
Jr. Prolongations Camaná N° 1014, Lima 01, Perú
(511) 998945990, jcancho@unfv.edu.pe

Appendix A: Questionnaire on Waste Management and the Circular Economy with Results

The purpose of this questionnaire is to gather information on current perceptions and practices in waste management in the district of San Juan de Lurigancho, as well as to assess knowledge and acceptance of technological solutions that promote the circular economy.

Instructions: Please answer the following questions honestly by marking the options you consider appropriate with a "×". Your answers will be confidential and will be used for research purposes only.

The following are the results obtained through the questionnaire applied to the residents of the San Juan de Lurigancho district.

Questionnaire on Waste Management and the Circular Economy with Results

Section	Question	Options	Results (%)
Demographic Data	Age:	☐ Under 18 years old	10
		☐ 18-30 years old	40
		☐ 31-50 years old	30
		☐ Más de 50 years old	20
	Genre:	☐ Male	50
		☐ Female	45
		☐ I'd rather not say	5
	Level of education:	☐ Primary	15
		☐ Secondary	25
		☐ Superior	40
		☐ Postgraduate	20
Knowledge and Perception of Waste Management	How would you rate waste management in your district?	☐ Very Deficient	15
		☐ Deficient	25
		☐ Adequate	30
		☐ Good	20
		☐ Very Good	10
	Are you aware of any recycling programs in your community?	☐ Yes	60
		☐ No	40
	How often do you separate your waste at home?	☐ Always	25
		☐ Often	30
		☐ Rarely	20
		☐ Never	25
Use of Technologies for Waste Management	Would you be willing to use a smart waste collection system (sensor-equipped bins) in your community?	☐ Yes	70
		☐ No	20
		☐ I'm not sure	10
	How important do you consider the implementation of technologies for waste management?	☐ Very important	55
		☐ Important	30
		☐ Not important	10
		☐ Nothing important	5
Circular Economy Practices	Are you familiar with the concept of a circular economy?	☐Yes	50
		☐ No	50
	Do you believe that adopting circular economy practices could benefit your community?	☐ Yes	65
		☐ No	25
		☐ I'm not sure	10
	What measures do you consider most effective to promote the circular economy in your community?	☐ Education and awareness	40
		☐ Economic incentives for recycling	25
		☐ Implementation of technologies	20
		☐ Improvements to waste management infrastructure	15
		☐ Others:	5

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20251030095053

Original scientific paper
Received: 2025-10-30, Accepted: 2025-11-28

Influence of the Process Parameters on the Mechanical Properties of the FDM 3D Printed Samples

Sara Havrlišan*, Martina Tokić, Štefanija Klarić, Branko Štrbac

Abstract: This research analyses the impact of infill density and layer thickness on the tensile strength of test samples fabricated using FDM 3D printing, with a particular focus on comparing samples produced with and without an adhesion layer. A two-factor factorial experimental design was employed, consisting of two levels and three replications. The infill density (20% and 50%) and layer thickness (0.1 mm and 0.2 mm) were defined as input variables, while tensile strength served as the output variable. Statistical methods, such as ANOVA, were used to analyse the experimental data and evaluate the significance of the individual factors and their interactions. The results revealed that for samples with an adhesion layer, the infill density, layer thickness, and their interaction had a statistically significant effect on tensile strength. In contrast, for samples without an adhesion layer, a significant effect was observed only for individual factors, with no significant interaction. The regression models exhibited high coefficients of determination (R^2) for both sample sets, confirming their strong reliability in predicting tensile strength. Furthermore, the t-test results indicated no statistically significant difference in tensile strength between samples manufactured with and without an adhesion layer.

Keywords: analysis of variance; FDM; regression model; tensile strength

1 INTRODUCTION

3D printing technologies enable the production of complex, geometrically demanding and individually tailored components that are difficult, and often impossible, to produce using conventional manufacturing methods. Thanks to the ability to create objects layer by layer, from digital 3D models, this technology has opened up completely new possibilities in various areas of human activity. Today, 3D printing is successfully applied in a wide range of sectors - from the manufacturing industry, medicine, architecture and construction, to the automotive, aerospace and space industries, but also in the fields of fashion, science and education.

A range of 3D printing technologies has been developed within the research and industrial community, including SLS (Selective Laser Sintering), SLA (Stereolithography) and FDM (Fused Deposition Modelling). FDM technology has become particularly popular and one of the most researched methods due to its simplicity, availability, lower cost and adaptability to a wide range of materials and applications [1-21].

The material selection for 3D printing depends on the purpose, functional, and mechanical and thermal properties requirements of the produced component. In FDM technology, PLA (polylactic acid), a biodegradable thermoplastic polymer that offers a good balance between ease of printing, precision, and environmental friendliness, is most often used. However, in order to improve the mechanical properties of the manufactured parts, scientific research is increasingly focusing on modifying existing materials through the application of various additives in composite materials, including carbon fibres [2, 8-11], as well as on the use of other materials such as ABS, PETG, etc. [16, 19].

At the same time, new printing technologies are being tested, with a detailed analysis of key mechanical properties - primarily tensile strength and bending test results and elastic modulus - which represent fundamental indicators of the quality and reliability of printed components [1-21].

Although significant progress has been made in the development of materials and the optimization of printing technologies in the last decade, the mechanical properties of 3D printed parts are still limited. The reason for this may relate to the number of parameters that simultaneously affect the printing process and are intertwined with each other, which makes it difficult to control and predict the results. The most commonly studied parameters include printing speed, infill density and pattern, layer thickness, layer width and height, filament and build plate temperature, print orientation, raster angle, nozzle temperature, and even model curvature [1-21].

In order to systematically examine the effects of the above parameters and find the optimal combination of printing parameters, the design of experiments method is often used in research. Among the most commonly used methods, the Taguchi method [1, 4-6, 15, 18] known for its efficiency in reducing the number of necessary experiments while maintaining high accuracy of results, and the response surface methodology (RSM) stand out [15, 19]. RSM allows a detailed analysis of interactions between parameters and modeling the relationship between input and output variables. Such approaches significantly contribute to the understanding of the FDM 3D printing process and enable the development of components with improved mechanical properties and optimised part characteristics.

The aim of this paper is to experimentally investigate the influence of selected FDM 3D printing process parameters, specifically the infill density and layer thickness, on the tensile strength of test samples made with and without an adhesion layer. The work applied a two-factor factorial design at two levels and three replications, where the infill density (20% and 50%) and the layer thickness (0.1 mm and 0.2 mm) were defined as input variables, while the tensile strength was observed as the output variable. The obtained experimental data were statistically analysed using the

analysis of variance (ANOVA) method to determine the significance of the individual factors and their interactions. Based on the results, regression models were created to predict the tensile strength values of samples with and without an adhesion layer. The influence of the adhesion layer itself on the tensile strength was also examined using the *t*-test.

2 EXPERIMENTAL PART

According to ISO 527 standard [22], twelve (12) test samples with an adhesion layer and twelve (12) test samples without an adhesion layer were made of GENERIC PLA material, using the FDM 3D printing process on the Wanhao Duplicator D12 300 Dual 3D printer, which is shown in Fig. 1.

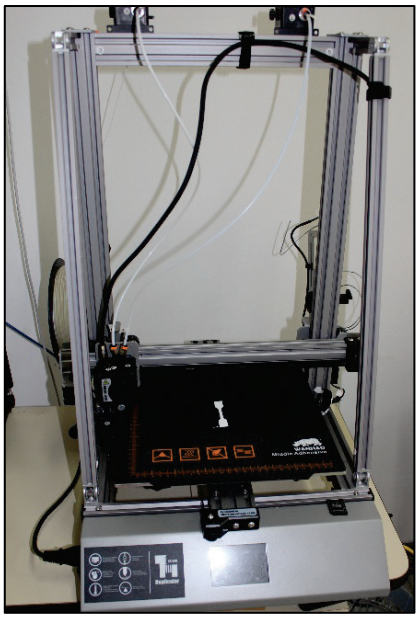


Figure 1 3D printer Wanhao Duplicator D12 300 Dual

A two-level factorial design with three replications was applied. When planning the experiment, the basic principles of experimental design were respected - replication and randomisation, and the principle of blocking was not applied. The input variables, or factors, are the infill density (20% and 50%) and the layer thickness (0,1 mm and 0,2 mm). The output variable, or response, is the tensile strength. In addition to the variables that changed (infill density and layer thickness), other factors were kept constant: nozzle temperature 210 °C, print speed 50 mm/s, worktable temperature 60 °C. The type of filling was Grid. The tensile strength was measured on a Shimadzu AGS-X device, as shown in Fig. 2. Test samples without an adhesion layer, prior to tensile strength testing, are shown in Fig. 3, while Fig. 4 shows the test samples after the test.

To investigate the influence of the adhesion layer on tensile strength, samples with an adhesion layer were fabricated. The adhesion layer is typically used to improve the bonding of the sample to the build platform and to prevent deformations during the printing process. The adhesion layer was made of the same material as the samples (GENERIC PLA material), with a thickness of 0,1 mm and a build plate adhesion type set to raft. An example of a test sample with an adhesion layer is shown in Fig. 5.



Figure 2 Shimadzu AGS-X tensile testing machine

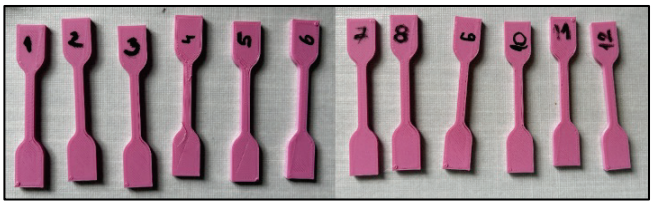


Figure 3 Test samples without an adhesion layer before tensile strength testing

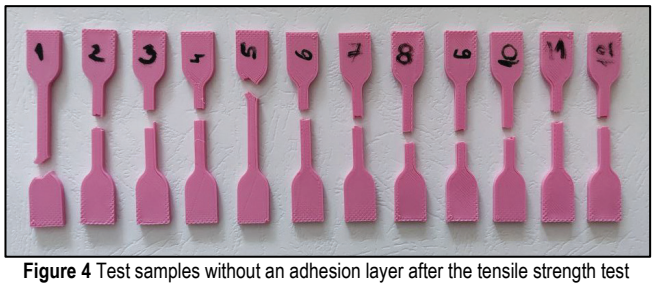


Figure 4 Test samples without an adhesion layer after the tensile strength test



Figure 5 Test sample with an adhesion layer

3 TENSILE STRENGTH ANALYSIS

Since the values of the infill density and the layer thickness were varied at two levels, a two-way ANOVA was performed to determine the significance of the influence of individual parameters and their interaction on the tensile strength of the samples. Additionally, a *t*-test was used to determine whether there was a statistically significant difference in the tensile strength values between samples made with and without an adhesion layer.

3.1 Tensile Strength Analysis of Samples with an Adhesion Layer

The design of the experiment and the measured tensile strength values of samples made with an adhesion layer are shown in Tab. 1. MS Excel was used for statistical analysis of data.

Table 1 Design of experiment and measured values of tensile strength for samples with an adhesion layer

Conventional layout	Trial no.	Infill density, %	Layer thickness, mm	Tensile strength 1, MPa
10	1	50	0,2	35,97
3	2	20	0,1	34,74
2	3	20	0,1	35,98
12	4	50	0,2	36,47
8	5	50	0,1	38,12
11	6	50	0,2	36,3
7	7	50	0,1	38,08
5	8	20	0,2	31,99
6	9	20	0,2	31,28
9	10	50	0,1	37,76
4	11	20	0,2	32,65
1	12	20	0,1	35,51

Tab. 2 shows the report from the Descriptive Statistics tool, which includes basic statistical parameters for samples with an adhesion layer.

The Data Analysis tool - ANOVA: Two-Factor with Replication was used for the analysis of variance. Tab. 3 shows the report of the ANOVA: Two-Factor with Replication tool for samples with an adhesion layer.

Table 2 Descriptive statistics tool reports for samples with an adhesion layer

Tensile strength 1, MPa	
Mean	35,40
Standard Error	0,67
Median	35,98
Mode	#N/D
Standard Deviation	2,32
Sample Variance	5,40
Kurtosis	–0,67
Skewness	–0,67
Range	6,84
Minimum	31,28
Maximum	38,12
Sum	424,85
Count	12

The average tensile strength value is 35,40 MPa (Mean), the median is 35,98 MPa (Median), the estimated standard deviation of the base set based on the sample data is 2.32 MPa (Standard Deviation), and the sample variance is 5,40 MPa (Sample Variance). The minimum tensile strength value is 31,28 MPa (Minimum), and the maximum value is 38,12 MPa (Maximum). From Tab. 1, it is evident that the minimum tensile strength value was achieved for a filling percentage of 20% and a layer thickness of 0,2 mm, while the maximum tensile strength value was achieved for a filling percentage of 50% and a layer thickness of 0,1 mm.

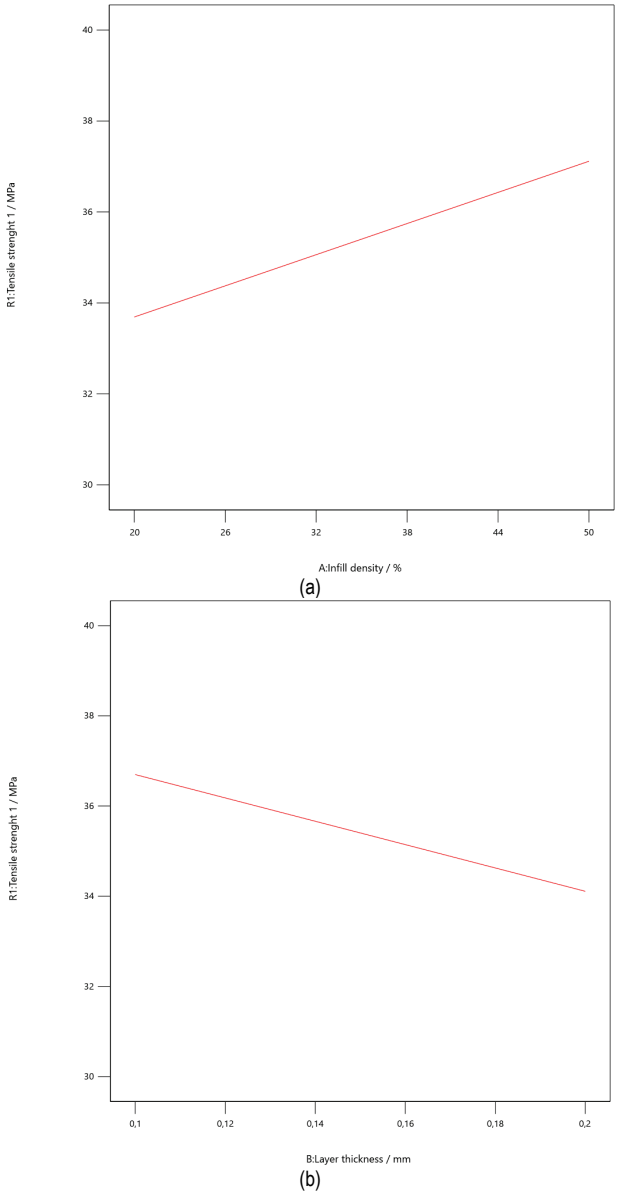


Figure 6 Corelation between tensile strenght 1 (samples with an adhesion layer) and (a) denisity infill; (b) layer thickness

Fig. 6 illustrates the correlation between tensile strength 1 (samples with an adhesion layer) and the parameters of infill density and layer thickness. A positive correlation is evident between tensile strength 1 and infill density, demonstrating that an increase in infill density enhances tensile strength (Fig. 6a). Conversely, a negative correlation is observed between tensile strength 1 and layer thickness, indicating that greater layer thickness diminishes tensile strength (Fig. 6b).

In order to investigate the influence of infill density and layer thickness factors on the response - tensile strength, it is necessary to compare the values of the calculated F ratios with the tabular values of the F variable. The value of the F ratio for the infill density factor is 145,89, and the corresponding tabular value of the F variable (F_{crit} value) is 5,32. The F-ratio is greater than the tabular value of the F variable, so it can be concluded that the change in the level of the infill density factor affects the change in tensile strength; the infill density is a significant factor. The same conclusion can be made based on the probability P-value for the calculated F-ratio. The probability for the calculated F ratio (P-value) is $2{,}04\times10^{-6}$, which is less than 0,05 (probability of a type I error α), so it can be concluded that the factor is significant, as the change in the level of the infill density factor affects the tensile strength.

Table 3 ANOVA tool report: Two-Factor with Replication for samples with an adhesion layer

Source of Variation	SS	df	MS	F	P-value	F_{crit}
Infill density	35,19	1	35,19	145,89	2E^{-06}	5,32
Layer thickness	20,10	1	20,10	83,32	1,7E^{-05}	5,32
Interaction	2,16	1	2,16	8,95	0,0173	5,32
Within	1,93	8	0,24			
Total	59,38	11				

The F-ratio value for the layer thickness factor is 83,32, and the corresponding table value of the F variable (F_{crit} value) is 5,32. The F-ratio is greater than the table value of the F variable, so it can be concluded that changing the level of the layer thickness factor also affects the change in tensile strength, i.e. the layer thickness is significant. The probability for the calculated F-ratio (P-value) is $1{,}67\times10^{-5}$, which is less than 0,05 (probability of a type I error α), so it can be concluded that the factor is significant; changing the level of the layer thickness factor affects tensile strength.

The F-ratio value for the interaction of the factors - infill density and layer thickness (Interaction) is 8,95 and the corresponding table value of the F variable (F_{crit} value) is 5,32. The F ratio is greater than the table value of the F variable, so it can be concluded that the interaction of factors A (infill density) and B (layer thickness) also affects the change in tensile strength; thus, the interaction of infill density and layer thickness is significant. The probability for the calculated F-ratio (P-value column) is 0,0173, which is less than 0,05 (probability of a type I error α), so it can be concluded that the interaction of the factors is significant, and the interaction of infill density and layer thickness affects tensile strength.

In order to obtain a model that will predict the dependence of tensile strength on the factors - infill density and layer thickness, it is necessary to conduct regression modelling. To obtain the regression model, the Regression tool in MS Excel was used.

For samples with an adhesion layer, the following values of regression coefficients were obtained: $\beta_0 = 38{,}26$; $\beta_1 = 0{,}029$, $\beta_2 = -45{,}678$, $\beta_{12} = 0{,}565$ based on which a regression model can be written, Eq. (1):

$$y = 38{,}26 + 0{,}029 \cdot x_1 - 45{,}678 \cdot x_2 + 0{,}565 \cdot x_1 \cdot x_2 \qquad (1)$$

where y is the response - tensile strength, x_1 is the factor - infill density, and x_2 is the factor - layer thickness, $x_1 x_2$ is the interaction between two factors - infill density and layer thickness. The regression model with real variables is, Eq. (2):

$$\begin{aligned} tensile\,strength = {} & 38{,}26 + 0{,}029 \cdot infill\,density - 45{,}678 \cdot \\ & \cdot layer\,thickness + 0{,}565 \cdot infill\,density \cdot layer\,thickness \end{aligned} \qquad (2)$$

The tensile strength will increase by 0,029 MPa for each unit increase in the percentage of filling, while it will decrease by 45,678 MPa for each unit increase in the layer thickness. For the obtained regression model, the coefficient of determination R^2 is 0,9675.

Tab. 4 presents the measured tensile strength 1 values, the corresponding predicted values (calculated according to Eq. (2)), and the residuals. The results demonstrate that the model exhibits minimal discrepancy between the actual and predicted tensile strength 1 values. The high coefficient of determination (R^2) further confirms that the model demonstrates strong predictive accuracy for tensile strength 1.

Table 4 Actual and predicted values, and residuals for tensile strength 1 (samples with an adhesion layer)

Trial no.	Infill density, %	Layer thickness, mm	Actual tensile strength 1, MPa	Predicted tensile strength 1, MPa	Residual, MPa
1	50	0,2	35,97	36,25	–0,28
2	20	0,1	34,74	35,41	–0,67
3	20	0,1	35,98	35,41	0,57
4	50	0,2	36,47	36,25	0,22
5	50	0,1	38,12	37,99	0,13
6	50	0,2	36,30	36,25	0,05
7	50	0,1	38,08	37,99	0,09
8	20	0,2	31,99	31,97	0,02
9	20	0,2	31,28	31,97	–0,69
10	50	0,1	37,76	37,99	–0,23
11	20	0,2	32,65	31,97	0,68
12	20	0,1	35,51	35,41	0,10

3.2 Tensile Strength Analysis of Samples without an Adhesion Layer

The experimental design and measured tensile strength values of samples made without an adhesion layer are shown in Tab. 5.

Tab. 6 shows the Descriptive Statistics tool report with basic statistical parameters for samples without an adhesion layer. The average tensile strength value is 36,28 MPa (Mean), the median is 36,25 MPa (Median), the estimated standard deviation of the base set based on the sample data is 2,63 MPa (Standard Deviation), and the sample variance is 6,90 MPa (Sample Variance). The minimum tensile strength

value is 32,65 MPa (Minimum), and the maximum value is 40,27 MPa (Maximum). From Tab. 5, it is evident that the minimum tensile strength value is achieved for a filling percentage of 20% and a layer thickness of 0,2 mm, while the maximum tensile strength value is achieved for a filling percentage of 50% and a layer thickness of 0,1 mm.

Table 5 Design of experiment and measured values of tensile strength for samples without an adhesion layer

Conventional layout	Trial no.	Infill density, %	Layer thickness, mm	Tensile strength 2, MPa
10	1	50	0,2	36,77
3	2	20	0,1	34,45
2	3	20	0,1	35,73
12	4	50	0,2	37,54
8	5	50	0,1	39,59
11	6	50	0,2	37,7
7	7	50	0,1	40,27
5	8	20	0,2	32,65
6	9	20	0,2	33,18
9	10	50	0,1	39,22
4	11	20	0,2	33,51
1	12	20	0,1	34,74

Table 6 Descriptive statistics tool reports for samples without an adhesion layer

Tensile strength 2, MPa	
Mean	36,28
Standard Error	0,76
Median	36,25
Mode	#N/D
Standard Deviation	2,63
Sample Variance	6,90
Kurtosis	–1,39
Skewness	0,13
Range	7,62
Minimum	32,65
Maximum	40,27
Sum	435,35
Count	12

Fig. 7 illustrates the correlation between tensile strength 2 (samples without an adhesion layer) and the parameters of infill density and layer thickness. As observed for tensile strength 1, a positive correlation is also present in this case between tensile strength 2 and infill density, with higher infill density resulting in increased tensile strength (Fig. 7a). Conversely, a negative correlation is evident between tensile strength 2 and layer thickness, with an increase from 0,1 mm to 0,2 mm leading to a reduction in tensile strength (Fig. 7b).

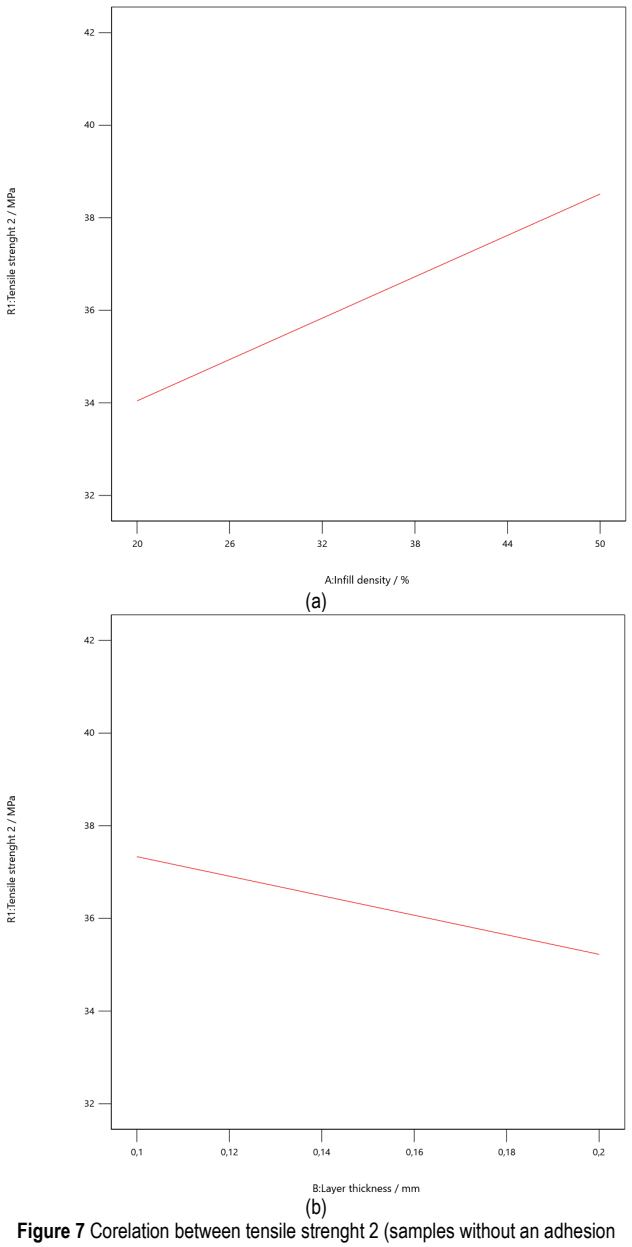


Figure 7 Corelation between tensile strenght 2 (samples without an adhesion layer) and (a) denisity infill; (b) layer thickness

Tab. 7 shows the ANOVA: Two-Factor with Replication tool report for samples without an adhesive layer.

Table 7 ANOVA tool report: Two-Factor with Replication for samples without an adhesion layer

Source of Variation	SS	df	MS	F	P-value	F_{crit}
Infill density	59,99	1	59,99	205,17	$5{,}51E^{-07}$	5,32
Layer thickness	13,33	1	13,33	45,61	0,000145	5,32
Interaction	0,19	1	0,19	0,63	0,4493	5,32
Within	2,34	8	0,29			
Total	75,85	11				

The F-ratio value for the infill density factor is 205,17 and the corresponding table value of the F variable (F_{crit} value) is 5,32. The F-ratio is greater than the table value of the F variable, indicating that changing the level of the infill density factor affects the change in tensile strength; therefore, the percentage of filling is a significant factor. The same conclusion can be drawn based on the probability P-value for the calculated F-ratio. The probability for the calculated F-ratio (P-value) is $5{,}51\times10^{-7}$, which is less than 0,05 (probability of a type I error α), so it can be concluded that the factor is significant, and changing the level of the infill density factor affects the tensile strength.

The F-ratio value for the layer thickness factor is 45,61, and the corresponding table value of the F variable (F_{crit}

value) is 5,32. The *F*-ratio is greater than the table value of the *F* variable, so it can be concluded that the change in the level of the layer thickness factor also affects the change in tensile strength, i.e. the layer thickness is significant. The probability for the calculated *F*-ratio (*P*-value) is 0,000145, which is less than 0,05 (probability of a type I error α), so it can be concluded that the factor is significant, i.e. the change in the level of the layer thickness factor affects tensile strength.

The *F*-ratio value for the interaction of the factors - infill density and layer thickness is 0,63, and the corresponding tabular value of the *F* variable (F_{crit} value) is 5,32. The *F*-ratio is smaller than the tabular value of the *F* variable, so it can be concluded that the interaction of the factor does not affect the change in tensile strength, - the interaction of infill density and layer thickness is not significant. The probability for the calculated *F*-ratio (*P*-value) is 0,4493, which is greater than 0,05 (probability of a type I error α), so it can be confirmed that the interaction of the factor is not significant.

For samples with an adhesion layer, regression modeling was also performed with the aim of obtaining a model that would predict the dependence of tensile strength on the infill density and layer thickness factors. For samples without an adhesion layer, the following values of regression coefficients were obtained: $\beta_0 = 34{,}225$; $\beta_1 = 0{,}149$ and $\beta_2 = -21{,}083$, based on which a regression model can be written, Eq. (3):

$$y = 34{,}225 + 0{,}149 \cdot x_1 - 21{,}083 \cdot x_2 \tag{3}$$

where y is the response - tensile strength, x_1 is the factor - infill density, and x_2 is the factor - layer thickness. The regression model with real variables is, Eq. (4):

$$\begin{aligned} &tensile\ strength = 34{,}225 + 0{,}149 \times infill\ density \\ &-21{,}083 \times layer\ thickness \end{aligned} \tag{4}$$

The tensile strength will increase by 0,149 MPa for each unit increase in the infill density, while it will decrease by 21,083 MPa for each unit increase in the layer thickness. For the obtained regression model, the coefficient of determination R^2 is 0,967.

Table 8 Actual and predicted values, and residuals for tensile strength 2 (samples without an adhesion layer)

Trial no.	Infill density, %	Layer thickness, mm	Actual tensile strength 2, MPa	Predicted tensile strength 2, MPa	Residual, MPa
1	50	0,2	36,77	37,46	–0,69
2	20	0,1	34,45	35,10	–0,65
3	20	0,1	35,73	35,10	0,63
4	50	0,2	37,54	37,46	0,08
5	50	0,1	39,59	39,57	0,02
6	50	0,2	37,70	37,46	0,24
7	50	0,1	40,27	39,57	0,70
8	20	0,2	32,65	32,99	–0,34
9	20	0,2	33,18	32,99	0,19
10	50	0,1	39,22	39,57	–0,35
11	20	0,2	33,51	32,99	0,52
12	20	0,1	34,74	35,10	–0,36

Tab. 8 presents the results of the measured tensile strength 2 values (samples without an adhesion layer), the predicted values (calculated according to Eq. (4)), and the residuals. The table shows that the error between the actual and predicted tensile strength 2 values is small, as the model obtained from Eq. (4) exhibits a high coefficient of determination ($R^2 = 0.967$), indicating that it can reliably predict tensile strength 2.

3.3 Tests of the Equality of Two Means - t-test

In order to examine whether there is a significant difference in the tensile strength values of samples with and without an adhesion layer, a hypothesis test on the equality of arithmetic means of the basic sets based on the samples (*t*-test) was performed. The testing was performed using the Data Analysis Tool of MS Excel.

Since the variances of the basis sets are unknown, it is necessary to test the hypothesis of their equality. To test the hypothesis of equality of variances of the basis sets, the null and alternative hypotheses are, Eq. (5), [24]:

$$\begin{aligned} &H_0: \sigma_1^2 = \sigma_2^2 \\ &H_1: \sigma_1^2 \neq \sigma_2^2 \end{aligned} \tag{5}$$

The alternative hypothesis is two-tailed, and therefore, it is impossible to use the *F*-Test Two-Sample for Variances tool test. To conclude on the equality of variances of the basic sets, the F_0 variable is used, which is calculated according to Eq. (6), [24]:

$$F_0 = \frac{S_1^2}{S_2^2} \tag{6}$$

where S_1^2 represents the variance of the first sample (test samples with an adhesion layer), and S_2^2 represents the variance of the second sample (test samples without an adhesion layer).

The variance of the first sample S_1^2 is 5,40, while the variance of the second sample S_2^2 is 6,90. The value of the F_0 variable is as follows, Eq. (7):

$$F_0 = \frac{S_1^2}{S_2^2} = \frac{5{,}40^2}{6{,}90^2} = 0{,}613 \tag{7}$$

It is necessary to check whether the calculated value of F_0 is within the limit values, Eq. (8), [24]:

$$F_{\alpha/2, n_1-1, n_2-1} < F_0 < F_{1-\alpha/2, n_1-1, n_2-1} \tag{8}$$

The limit values of the F_0 variable are 3,474 and 0,288. The calculated value of the F_0 variable is within the limit values, and it can be concluded that the variances of the basic sets from which the samples were taken are equal. Based on the conclusion reached, the hypothesis of equality of the

arithmetic means of the basic sets can be tested using the *t-Test: Two-Sample Assuming Equal Variances tool*. The null and alternative hypotheses are, Eq. (9), [24]:

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2 \quad (9)$$

Tab. 9 shows the report of the *t*-test: Two-Sample Assuming Equal Variances tool.

Table 9 Report of the t-Test: Two-Sample Assuming Equal Variances tool

t-Test: Two-Sample Assuming Equal Variances		
	Tensile Strength 1, MPa	Tensile Strength 2, MPa
Mean	35,40	36,28
Variance	5,40	6,90
Observations	12	12
Pooled Variance	6,15	
Hypothesized Mean Difference	0	
df	22	
t Stat	–0,86	
P(*T* <= *t*) one-tail	0,199	
t Critical one-tail	1,72	
P(*T* <= *t*) two-tail	0,39	
t Critical two-tail	2,07	

Since the alternative hypothesis is two-tailed, it is necessary to compare the value of the calculated variable *t Stat* with the threshold value *t Critical two-tail*. The variable *t Stat* is –0,86, and the threshold value *t Critical two-tail* is 2,07. Since the value *t Stat* is less than *t Critical two-tail*, the alternative hypothesis is rejected, i.e. the null hypothesis of the equality of the arithmetic means of the basic sets is accepted. The same conclusion can be drawn based on the probability value P ($T \leq t$) two-tail. Since the probability is 0,39, which is greater than 0,05 (the probability of a type I error α), the null hypothesis is accepted, and it is concluded that there is no significant difference in the values of the tensile strengths of samples with and without an adhesion layer.

4 CONCLUSION

In this paper, the influence of the infill density and layer thickness on the tensile strength of samples manufactured by FDM 3D printing was experimentally investigated, with the aim of developing regression models for predicting mechanical properties. The research was conducted using a factorial experimental design with two levels and three replications, and the test samples were manufactured with and without an adhesion layer. The obtained results were statistically analysed using the analysis of variance (ANOVA) method in MS Excel.

The analysis of samples with an adhesion layer revealed that the infill density, the thickness of the layer and their interaction have a significant impact on the tensile strength, while for samples without an adhesion layer a significant impact was confirmed for both individual factors, but not for their interaction. In both cases, regression models with high coefficients of determination were created ($R^2 = 0.9675$ for samples with an adhesion layer and $R^2 = 0.967$ for samples without an adhesion layer), which confirmed their high accuracy in predicting the tensile strength.

Using a *t*-test, it was determined that there is no statistically significant difference in tensile strength values between samples made with and without an adhesion layer ($p = 0.39 > 0.05$), indicating that the adhesion layer does not significantly affect the mechanical properties. Therefore, it can be concluded that making samples without an adhesion layer can save time and material consumption without negatively affecting the tensile strength of the printed samples, provided that the printing process and product geometry are not affected.

Acknowledgements

This research paper was funded by the University of Slavonski Brod through the institutional research project Smartmod4IND45, financed by the European Union - NextGenerationEU. The views and opinions expressed in this paper are those of the author and do not necessarily reflect the official position of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

5 REFERENCES

[1] Ekşi, S., & Karakaya, C. (2025). Effects of process parameters on tensile properties of 3D-printed PLA parts fabricated with the FDM method. *Polymers*, *17*(14), 1934. https://doi.org/10.3390/polym17141934

[2] Hamoud, M., Elshalakany, A. B., Gamil, M., & Mohamed, H. (2024). Investigating the influence of 3D printing parameters on the mechanical characteristics of FDM fabricated (PLA/Cu) composite material. *The International Journal of Advanced Manufacturing Technology, 134*, 3769-3785. https://doi.org/10.1007/s00170-024-14313-0

[3] Gajjar, T., Yang, R. (Chunhui), Ye, L., & Zhang, Y. X. (2025). Effects of key process parameters on tensile properties and interlayer bonding behavior of 3D printed PLA using fused filament fabrication. *Progress in Additive Manufacturing, 10*, 1261-1280. https://doi.org/10.1007/s40964-024-00704-y

[4] Fidan Bilir Kilinc, F., Turkoglu, T., Guler, S., & Cagri Kilinc, A. (2025). Optimization of 3Dprinting parameters for enhanced tensile properties in continuous carbon fiber reinforced PLA composites. *Materials Research Express*, *12*(2025), 045302. https://doi.org/10.1088/2053-1591/adc5cc

[5] Layeb, N., Barhoumi, N., Oldal, I., & Keppler, I. (2025). Improving the strength properties of PLA acetabular liners by optimizing FDM 3D printing: Taguchi approach and finite element analysis validation. *The International Journal of Advanced Manufacturing Technology*, *137*, 2649-2664. https://doi.org/10.1007/s00170-025-15314-3

[6] Megersa, G. K., Sitek, W., Nowak, A. J., & Tomašić, N. (2024). Investigation of the influence of fused deposition modeling 3D printing process parameters on tensile properties of polylactic acid parts using the Taguchi method. *Materials, 17*(23), 5951. https://doi.org/10.3390/ma17235951

[7] Khan, H., Siddiqi, M. ur R., Saher, S., Muhammad, R., & Saad, M. (2024). Tensile properties of 3D-printed PLA prismatic cellular structures: an experimental investigation. *The*

International Journal of Advanced Manufacturing Technology, *134*, 4399-4410. https://doi.org/10.1007/s00170-024-14343-8

[8] Mirza, F., Shenoy, S. B., Nunna, S., Kini, C. R., & Creighton, C. (2024). Effect of material extrusion process parameters on tensile performance of pristine and discontinuous fibre reinforced PLA composites: A review. *Progress in Additive Manufacturing*, *10*, 3251-3265. https://doi.org/10.1007/s40964-024-00825-4

[9] Blake, B., Mendenhall, R., & Eslami, B. (2025). Balancing strength and flexibility: Mechanical characterization of carbon fiber-reinforced PLA composites in FDM 3D printing. *Journal of Manufacturing and Materials Processing*, *9*(9), 288. https://doi.org/10.3390/jmmp9090288

[10] Ropovik, R., Peti, D., & Hatala, M. (2025). Strength analyses of FDM products from additively strengthened PLA materials. *MM Science Journal*, 8512-8519. https://doi.org/10.17973/MMSJ.2025_09_2025085

[11] Singh, D., & Verma, A. (2025). Optimization of process parameters for maximizing tensile strength in 3D-printed ASTM specimens: A comparative study of Black PLA and Carbon PLA. *ITM Web of Conferences, 74*, 03010. https://doi.org/10.1051/itmconf/20257403010

[12] Lee, N., Hamid, R. A., Maidin, S., Hajar, N., & Ito, T. (2024). Effect of process parameters on tensile strength of 3D printed PLA parts. *Journal of Advanced Manufacturing Technology (JAMT)*, *18*(3).

[13] Efa, D. A., & Surya, D. (2025). Optimization of design parameters and 3D-printing process using RSM and optimization algorithms. *Journal of Manufacturing and Materials Processing*, *26*, 100702. https://doi.org/10.1016/j.jmmp.2025.470

[14] Tunalioglu, M. S. (2025). The investigation of the working conditions on 3D printed helical gears. *Tehnicki vjesnik - Technical Gazette*, *32*(2), 536-541. https://doi.org/10.17559/TV-20241008002042

[15] Saefudin, S., Raharjo, S., Yustar Afif, I., Syarifudin, Purnomo, Rusydi, M. O., Kuzmin, A., Subri, M., & Pujianto, M. E. (2025). A Review of Factors Affecting the Mechanical Performance of PLA in FDM 3D Printing. *Advance Sustainable Science, Engineering and Technology*, *7*(2), 0250208. https://doi.org/10.26877/chs1gc62

[16] Lerher, T., Benčak, P., Ebrahim, S., Motaln, M., & Hercog, D. (2025). Design and implementation of soft robotic gripper using 3D printing technology. *Tehnički glasnik - Technical Journal*, *19*(SI), 112-117. https://doi.org/10.31803/tg-20250411222835

[17] Pyka, D., Słowiński, J. J., Kurzawa, A., Roszak, M., Stachowicz, M., Kazimierczak, M., Stępczak, M., & Grygier, D. (2024). Research on Basic Properties of Polymers for Fused Deposition Modelling Technology. *Applied Sciences*, *14*(23), 11151. https://doi.org/10.3390/app142311151

[18] Durović, S., Lazarević, D., Milovanović, V., Mišić, M., & Šarkočević, Ž. (2024). The influence of the number of layers and the raster angle on the mechanical properties of 3D printed materials. *Tehnički vjesnik - Technical Gazette*, *31*(6), 1892-1897. https://doi.org/10.17559/TV-20240325001430

[19] Ben hadj Hassine, S., Chatti, S., Louhichi, B., & Seibi, A. (2024). Experimental study of the tensile behavior of structures obtained by FDM 3D printing process. *Polymers*, *16(*11), 1562. https://doi.org/10.3390/polym16111562

[20] Karamanlı, İ. A., & Tahnal, K. (2024). Optimization of printing parameters of PLA and ABS produced by FFF. *Journal of Materials and Mechatronics: A*, *5*(2), 286-302. https://doi.org/10.55546/jmm.1566700

[21] Camargo, J. C., Machado, Á. R., Almeida, E. C., & Silva, E. F. M. S. (2019). Mechanical properties of PLA-graphene filament for FDM 3D printing. *The International Journal of Advanced Manufacturing Technology*, *103*(7-8), 2423-2443. https://doi.org/10.1007/s00170-019-03532-5

[22] ISO 527-1:2019. *Plastics - Determination of tensile properties - Part 1: General principles*. International Organization for Standardization.

[23] Tokić, M. (2024). *Software support in statistical analysis*. Master's thesis, University of Slavonski Brod, Mechanical Engineering Faculty.

[24] Montgomery, D. C. (2013). *Design and Analysis of Experiments.* John Wiley & Sons: Hoboken, USA, 730.

Authors' contacts:

Sara Havrlišan, PhD, Assistant Professor
(Corresponding author)
University of Slavonski Brod,
Mechanical Engineering Faculty in Slavonski Brod,
Trg I. B. Mažuranić 2, 35000 Slavonski Brod, Croatia
E-mail: shavrlisan@unisb.hr

Martina Tokić, Student
University of Slavonski Brod,
Mechanical Engineering Faculty in Slavonski Brod,
Trg I. B. Mažuranić 2, 35000 Slavonski Brod, Croatia

Štefanija Klarić, PhD, Professor
Charles Darwin University,
Ellengowan Drive, Casuarina,
NT 0810, Australia
E-mail: Stefanija.klaric@cdu.edu.au

Branko Štrbac, PhD, Associate Professor
University of Novi Sad,
Faculty of Technical Sciences,
Trg Dositeja Obradovića, 21000 Novi Sad, Serbia
E-mail: strbacb@uns.ac.rs

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20251108135839

Original scientific paper
Received: 2025-11-08, Accepted: 2025-12-08

Experimental Performance Analysis of PCM and Fin-Integrated Solar Stills under Varying Water Depths

Kailas Deore*, Nilesh Salunke, Sudhakar Pawar, Yogeshkumar Pathak, Juber Khatik

Abstract: Limited availability of freshwater has increased the demand of low-cost and sustainable desalination technologies. Basin-type solar stills are a simple and cost-effective solution with a limitation of low thermal efficiency and low daily productivity. To overcome these issues, the current study experimentally analyzes how solar stills can be improved in terms of performance by incorporating phase change materials (PCM) and heat transfer fins. Three designs were constructed and experimented with basin water depths of 2 cm, 3 cm and 4 cm; the conventional solar still (CSS), a PCM-integrated still and a finned PCM-integrated still (FPCM). Thermocouples, pyranometer and graduated collection were systematically instrumented to measure temperature profiles, solar irradiance and yield of distillate. The findings indicate that the FPCM setup was always the most successful, and the daily productivity was about 325 mL/m²/day at 2 cm depth, which is about 40% better than CSS. The PCM-based still yielded intermediate enhancement, with 200-210 mL/m 2/day at 3 cm depth, which is indicative of the correspondence between water mass and PCM charging/discharging performance. CSS had a thermal efficiency of 1.623% versus FPCM of 2.534% and CSS had productive hours of approximately 11 h versus 12 h in FPCM. The peak hourly yield was 50-55 mL/h in FPCM and 38 mL/h in CSS. These results prove that PCM integration improves thermal storage and fins further speed up heat transfer resulting in high-quality distillate yield and long operation. The research confirms that FPCM solar stills is a feasible, inexpensive method of freshwater production in water-deprived areas.

Keywords: desalination; experimental study; freshwater productivity; heat transfer fins; Phase Change Material (PCM); solar still; thermal efficiency; water depth

1 INTRODUCTION

The issue of freshwater scarcity has become one of the most significant worldwide challenges due to the rapid increase in population, industrialization, urbanization, and the complex effects of climate change, which has put an intensive strain on the available water resources, and threatened the lives of billions of people globally [1, 2]. Desalination technologies are the critical solutions to the growing need of drinking water in arid and semi-arid areas by transforming salty or brackish water into drinking supplies [3,4]. However, the traditional desalination systems are generally described as being energy-consuming, significant environmental footprint, and high cost economically, thus, fuelling the development of sustainable alternatives. Solar desalination, especially the implementation of solar stills, is a relatively inexpensive, ecologically neutral, and decentralised approach to freshwater production that makes it especially suitable to rural and remote populations with a large solar resource base [5]. Renewable solar power is incorporated to enable evaporation and condensation processes in these systems, which forms a feasible course of available water purification.

Today, solar distillation is an environmentally friendly and inexpensive method. The main drawback is that solar stills are extremely inefficient heaters and produce very small amounts of water per day [6]. Traditional solar stills are based on the basic principle of the evaporation-condensation cycle, according to which solar radiation warms up saline or brackish water that is in a basin, and thus, leads to its evaporation. The obtained vapour is then condensed onto the colder side of a glass cover to produce drinkable water, which is then stored separately [7]. This design has been praised due to its simplicity, cost-efficiency, and low maintenance, which make it especially appealing in decentralized applications in remote or resource-constrained environments [8, 9]. However, despite these benefits, traditional solar stills face significant technical challenges. The latter include low productivity by default, high heat losses, reduced evaporation rates, and high reliance on the daytime solar irradiance, all of which severely limit the freshwater yield [10, 11]. In particular, the limited temperature difference between the evaporating and condensing surfaces, inefficient use of heat, and the lack of thermal storage often lead to relatively small daily yields, which are usually less than 4 L/m^2/day, which explains the urgency of technological advances to improve efficiency [12].

The use of phase change materials (PCMs) to improve the performance of solar stills has been shown to capture thermal energy during the brightest sunshine hours and subsequently release it during the dark hours to prolong operation beyond the daylight hours [13]. Fins are also added to the system to enhance heat transfer between the absorber and the water, thus increasing the rate of evaporation by increasing the heat transfer area and the even distribution of heat [14]. The synergistic effect of PCMs and fins is even more productive, because the fins increase the effective thermal conductivity of the PCM to offset its low heat transfer rates, and increases daytime and nighttime freshwater yield [15, 16]. The depth of water in the basin is a vital working factor: the lower the depth, the higher the rate of evaporation, but the deeper the layer, the lower the temperature gradient and, consequently, the performance efficiency [17, 18]. Recent studies have compared conventional, PCM-enhanced, and fin-augmented stills and have found that there are considerable improvements in efficiency, productivity, and sustainability [19]. However, there is still a research gap regarding systematic experimental comparisons between PCM-only, fin-only, and fin-plus-PCM solar stills, especially when the water depth and climatic conditions vary, which hinders a thorough comprehension of the integrated performance capabilities of these systems [20, 21].

The recent developments in the solar still technology have majorly focused on increasing the productivity, thermal efficiency and cost effectiveness by incorporating phase

change materials (PCMs), nanoparticles, geometric alterations and eco-friendly materials. Scientists have explored numerous options to address the fundamental shortcomings of traditional stills, i.e. low yields, heat loss and intermittency. Ahmed et al. [22] showed that the introduction of 40 mm PCM encapsulated cylindrical fins into pyramid solar stills optimized the water depth, enhancing the productivity by 49.9% and efficiency by 28.13% to 36.42%, and Jahapanah et al. [23] showed that the addition of 6 kg of low-temperature PCM into single-slope solar stills increased productivity by 30.3% and efficiency by 28.13% to The oval and tubular solar still with cover cooling and PCM proposed by Aly et al. [24] reached a 32.42% productivity increase and a peak of 6.78 L/m^2/day. Likewise, Afolabi et al. [25] created a solar still with a double-slope and microencapsulated PCM-TES, providing an impressive 105% productivity gain, three-hour extension, and a short payback period of 0.8 years.

Geometric improvements have been also critical as shown by Toosi et al. [26] who showed that the addition of an external condenser and phase-change material (PCM) in a stepped solar still increased the productivity by 104%. At the same time, Dawood et al. [27] have found that a solar dish added to a stepped-slope PCM-containing solar still produced a maximum of 13.9 kg/day of freshwater at the optimal flow rate of 550 mL/min. Hussen et al. [28] established that the pyramid-shaped solar still with a raised basin and PCM -Ag nanoparticles yielded the highest productivity and thermal efficiency increases of 116.4 and 51.3% respectively, and reduced CO_2 emissions and water expenses. Simultaneous studies of PCM composition and structure have shown good results. Surapararaj and Natarajan [15] proposed a single-basin solar still with a pin-finned absorber in a paraffin wax bed, which led to the increase of the yield by 24.26% and carbon credits by the same percentage. According to Sahu et al. [29], the use of RT -58 PCM increased productivity by 46% and economic performance. Agrawal and Singh [30] showed that the use of a binary eutectic PCM in the solar stills with two slopes increased the energy efficiency to 22.21% to 30.42%, which was supported by computational fluid dynamic simulations. Murali et al. [31] also extended this strategy by incorporating Al_2O_3 nanoparticle-enhanced PCM, and the productivity and efficiency were increased by 60.37 and 68.29%, respectively.

Material innovations have contributed to the advancement of technology. Goshayeshi et al. [32] found that 0.5% of graphene oxide in paraffin in a semicircular absorber was the most effective, and Shalaby et al. [33] found that a paraffin-CuO nanocomposite latent heat storage system can increase productivity by 4.54% and has significant potential to reduce costs at scale. Elmghari et al. [34] highlighted the significance of optimization of the PCM mass where a single-sloped configuration with 2kg of PCM still provided a 27.7% productivity gain and better energy and exergy efficiency compared to configurations with higher PCM masses. Simultaneously, Boopalan et al. [35] incorporated paraffin-wax PCM into copper tubes, and thus, increased clean-water generation by 11% during the night. Lastly, Natarajan et al. [36] sought environmentally friendly improvements by using sawdust and rice straw, which increased the rate of evaporation and condensation and resulted in an impressive 62.88% productivity increase, as well as the lowest cost per litre and the shortest payback period of the tested configurations.

The recent developments in solar desalination technology have focused on improving thermal performance and fresh water output by using PCMs, advanced fin geometry and hybrid heat transfer techniques. Alqsaier et al. [37] showed that a modified hemispherical solar still with a PCM-Ag composite, rear reflectors, and fan-assisted heat transfer had a productivity of up to 172% higher compared to a standard design, producing 10.5 L/m^2-day of freshwater. This observation highlights the effectiveness of thermal and geometric additions. Equally, Bady et al. [38] recorded that conical solar distillers with hollow copper fins loaded with PCM achieved a freshwater production of 7.50 L/m^2-day and an efficiency of 80.3%, which was 54.6% higher than conventional designs and 550% higher than night-time productivity. Prasad et al. [39] experimentally validated that a single-sloped solar still with a thermal energy storage system, v-corrugated and parallel plate fins, achieved a productivity increase of 77.34% and a yield of 4.54 L/m^2-day and also minimized carbon emissions and water production costs. Moreover, Saad et al. [40] established that the integration of longitudinal fins in a PCM layer enhanced thermal efficiency by 45.12% and freshwater yield by 51.13% in addition to a 7.08% decrease in the cost of production, which confirmed the feasibility and cost-efficiency of this type of hybrid designs.

Despite much research on basin-type solar stills, the majority of modern research is limited to the performance modelling of a simulation or isolated analyses of individual additions like PCMs or fin arrays. Few studies have studied the combined effects of PCM addition and fin arrays under different basin water depths and most of the literature published has not touched the practical implications of the water depth, PCM interaction on distillate yield, thermal efficiency and productive hours. Furthermore, the inconsistencies in the literature, especially the work of PCM in the shallow waters, underscore the need to systematically validate the experiment. In that regard, the current paper conducts an experimental evaluation of conventional, PCM-integrated, and finned PCM-integrated solar stills at three different water depths of the basin (2 cm, 3 cm, and 4 cm). The research aim is to measure the increases in the daily productivity, thermal efficiency, productive hours, peak hourly yield, and realistic operational variability during the experiment. Finally, it is hoped that the practical performance advantages of PCM and fin-enhanced solar stills will be established, and experimental evidence that will support their use as low-cost, sustainable freshwater generation systems in water-scarce areas will be provided.

2 EXPERIMENTAL METHODOLOGY AND DATA ANALYSIS

The research design of the study is a controlled experimental one, supplemented with systematic analytical evaluation and data collection procedures. The study includes field tests of three solar still designs, namely CSS, PCM and FPCM, which are tested in strictly controlled standardized environmental conditions to make them comparable. Calibrated Type-K thermocouples, a digital data acquisition (DAQ) system and a pyranometer were used to measure

temperature and solar irradiance. The distillate yields were measured using graduated measuring containers and load cells. Further data analysis and visualization were performed in Microsoft Excel and MATLAB, thus, making it easier to not only validate the results but also create figures that can be interpreted.

The experiment and the test procedures were carried out at the rooftop of the SSBT's College of Engineering and Technology, Jalgaon, India.

2.1 Experimental Test Setup

The experimental study was carried out to determine the performance of a conventional single-slope solar still (CSS) and a phase change material (PCM)-integrated solar still and a finned PCM (FPCM)-integrated solar still under the same operating conditions. The test system was configured to be very similar to the conditions that were taken into account in the previous MATLAB-based numerical analysis, yet, enough instrumentation was installed in order to measure the thermal and productivity parameters that were to be validated.

2.1.1 Solar Still Configuration

The solar stills were made with the same geometric parameters so that all the test cases were uniform. Each still was made of a square basin whose effective base area was 1.0 × 1.0 m and insulated sidewalls to reduce heat losses. The height of the front wall was kept at 100 mm and the height of the rear wall at 525 mm to create a cover angle of about 23 o to allow the condensate to run off into the collection channel. The top of each unit was covered with transparent glass covers of 4 mm thickness with silicone gaskets to make them airtight. The inner basin plate was made of galvanized iron which was painted black to ensure that the plate was as absorptive as possible to the solar radiation. A sloped channel was used to collect the condensate at the lower edge of the glass in graduated measuring jars.

Three configurations were investigated:

CSS (Conventional Solar Still): 1) PCM was not added, saline water was kept at 2 cm, 3 cm and 4 cm depth. 2) PCM-integrated Still: A sheet of paraffin wax-based PCM (melting range 35-45 °C, 20 mm thick) was laid under the basin liner. The PCM was put in cylindrical aluminum capsules to prevent leakage and to store heat during peak hours. 3) Finned PCM-integrated Still: The PCM set up was augmented with copper fins of 40 mm and 1 mm thickness that were placed with a density of around 10 fins per basin width as depicted in Fig. 1(b). The fins were attached to the basin liner and embedded partially in the PCM to provide better conduction routes and increase the charging/ discharging rate.

The schematic of the experimental setup of the finned PCM-integrated solar still is as shown in Fig. 1(a). The design is made on the basis of single slope basin-type design, which entails a transparent glass cover, a blackened basin liner and side insulation to reduce the heat losses. The radiation that enters the solar system diffuses through the transparent cover and is absorbed by the basin liner where it warms the layer of water.

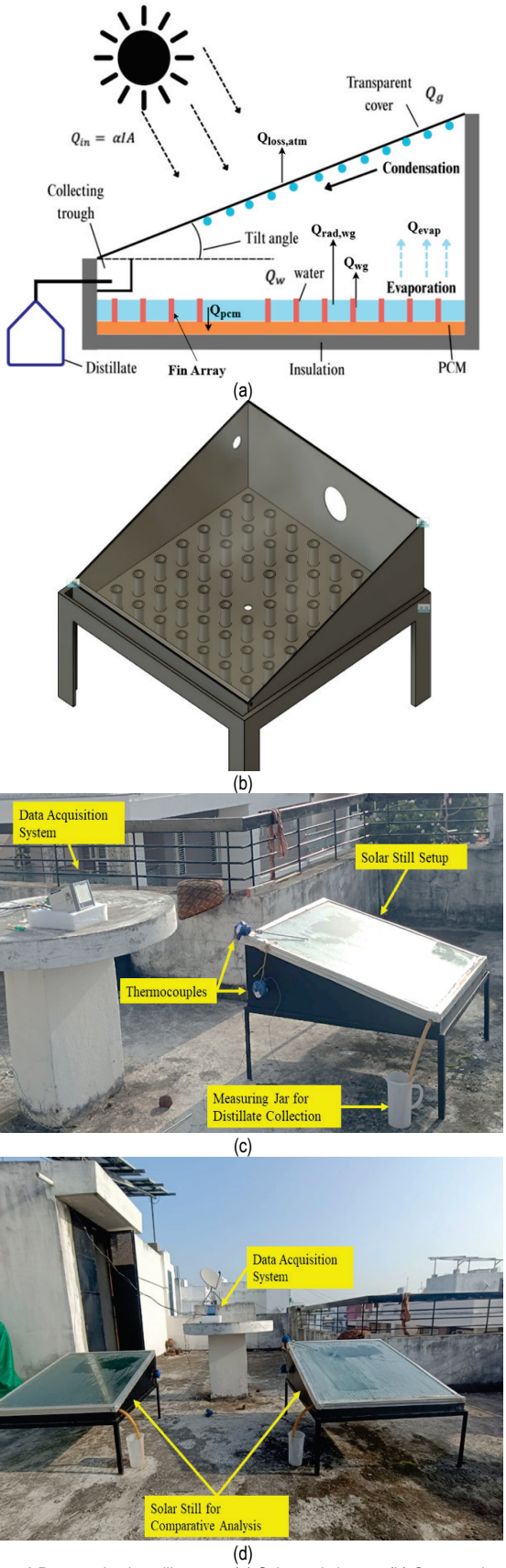


Figure 1 Proposed solar still system (a) Schematic layout, (b) Constructional features of finned solar still system, (c) Experimental test setup, and (d) Solar still system for comparative analysis

The basin has a fin array, which is connected to the phase change material (PCM) layer that is below the basin surface. The fins serve as thermal conductors, and they conduct heat quickly through the basin liner to the PCM, thus speeding up the charging process. A fraction of the absorbed heat is stored in the PCM in the form of latent heat (Q_{pcm}) during peak radiation hours and the remaining is used to raise the temperature of the water (Q_w). This accumulated energy is discharged slowly when the sun is not shining, and it maintains the water temperature, and evaporation. Water evaporates out of the basin at a rate that is dictated by the amount of heat taken in (Q_{evap}). The vapor of water is lifted up and condenses on the inner surface of the transparent glass cover which is lower in temperature because of the convectional cooling of the ambient air. The condensate moves down the inclined glass cover, and is collected in a trough, which is then directed to a measuring jar as distillate output.

The system is prone to multiple heat loss processes, such as radiation losses ($Q_{rad,w-g}$ of water to glass, evaporative losses (Q_{evap}, and convective/sky losses ($Q_{loss,atm}$) The PCM and fins inclusion minimizes the aggregate effect of these heat losses by increasing thermal storage capacity and guaranteeing efficient use of incident solar energy, which ultimately enhances freshwater productivity cumulative with respect to a traditional still. Fig. 1(c) and Fig. 1(d) shows experimental test setup and solar still systems for comparison under similar ambient conditions.

2.1.2 Instrumentation and Sensors

To quantify the thermal and productivity performance, multiple thermocouples were deployed at critical locations. Type-K thermocouples with an accuracy of ±0.5 C were selected, connected to a digital data acquisition (DAQ) system. The following sensor distribution was employed: Ambient Temperature (1 sensor): Shielded thermocouple placed in shade, 1.5 m above ground level. Water Temperature (2 sensors): One thermocouple placed at mid-depth and another 10 - 15 mm below the surface to capture evaporation dynamics. Basin Plate Temperature (1 sensor): Fixed to the underside of the basin plate at the geometric center with thermal paste. Glass Cover Temperature (1 sensor): Attached to the inner surface of the glass cover, midway along the slope, carefully avoiding condensate flow paths. PCM Temperatures (2 sensors): One at the core of the PCM encapsulation to monitor melting/solidification behavior and one at the PCM surface/fin base to capture heat transfer interface dynamics.

In total, five thermocouples were used for CSS, and seven for PCM and FPCM configurations. All sensors were logged at one-minute intervals using the DAQ system. Calibration of thermocouples was carried out prior to testing using ice bath (0 °C) and warm water (50 °C) benchmarks to minimize systematic error. A pyranometer was installed adjacent to the stills to measure solar irradiance (W), and a digital thermometer recorded the ambient air temperature to cross-check DAQ data. Hourly distillate yield was measured manually by recording the volume collected in graduated measuring jars. For continuous monitoring, in some trials, the jars were mounted on an electronic load cell (±1 g resolution), enabling automated recording of distillate mass. Electrical conductivity (EC) and total dissolved solids (TDS) were measured for both feed water and distillate samples using a handheld EC meter to confirm desalination quality. Tab. 1 provides technical specifications of instruments used in the experimental setup.

Table 1 Specifications of instruments used in the experimental setup

Instrument / Sensor	Measurement parameter	Accuracy / Resolution	Deployment location / Notes
Type-K Thermocouples	Temperature (ambient, water, basin plate, glass, PCM)	±0.5 °C	Ambient (1.5 m height), water (mid-depth & near surface), basin plate, glass, PCM core & surface
Data Acquisition System (DAQ)	Temperature logging	1-minute interval	Connected to all thermocouples for continuous data logging
Pyranometer	Solar irradiance	±3%	Positioned adjacent to stills in unobstructed sunlight
Digital Thermometer	Ambient temperature	±0.5 °C	Used to cross-verify ambient readings from thermocouples
Graduated Measuring Jars	Hourly distillate volume	±1 mL	Used for manual measurement of distillate collection
Load Cell	Distillate mass (automated)	±1 g	Jar mounted on load cell for real-time condensate mass measurement
Handheld EC Meter	Electrical conductivity (EC) & TDS	±2% (typical)	Used to assess desalination quality of feed water and distillate

2.2 Experimental Procedure

The experimental study was done on the rooftop laboratory of the institute so as to have unhindered exposure to the sun. All the tests were conducted in natural climatic conditions in summer when the chances of having clear sky days were high hence reducing the variability caused by cloud cover. Every experimental run was initiated at the beginning of the morning to coincide with the sunrise and ended at the end of the evening, so that both daytime charging and late-evening discharging response of the PCM could be observed. Each time it was run, the solar still basin was cleaned and the glass cover washed with distilled water to avoid any deposition of dust or salt that would change transmissivity. Saline water of known concentration (approximately 35%) was then added to the stills up to the desired depth of water (2, 3 or 4 cm depending on the case). A calibrated depth gauge and a tolerance of ±1 mm was used to measure the fill level to ensure consistency between trials. In PCM and finned PCM (FPCM) systems, the PCM modules were placed under the basin liner and the fins were checked to make sure that they were well bonded to the PCM and liner to prevent thermal resistance at the contact point. To ensure that melting properties are well captured during the day, the PCM was ensured to be in solid state at the beginning of the test. The experimental time table was systematic test matrix.

All Tier-1 cases, including CSS, PCM and FPCM at three different depths, were tested twice in clear-sky conditions to eliminate the effects of diurnal variations and to ensure the consistency of the results. To balance the exposure of each order of run to slightly different weather patterns, a Latin square rotation was used to allocate the order of run. The hourly readings of distillate yield were made during operation by recording the volume of condensate in graduated jars. In the case of the runs in which load cells were employed, the mass of distillate was automatically logged reducing the amount of manual labor. The sky condition was also taken at every hour (clear, partially cloudy, hazy) to contextualise the changes in irradiance. The cumulative distillate yield was measured and taken. The quality of the distillate was determined by the electrical conductivity (EC) and total dissolved solids (TDS) using a handheld meter. At the beginning of the day, the salinity baseline was also checked using the feed water. Every instrument was turned off, sensor was checked, and PCM modules were examined to make sure that there was no leakage or degradation. This was done systematically up until all the test cases were done.

2.3 Data Collection and Post-processing

The thermocouples were recorded to a multi-channel data acquisition (DAQ) system at a one-minute time interval, providing high-resolution records of temperature transients. The channels were ambient temperature, water temperature at mid-depth and near-surface, basin plate temperature, inner glass cover temperature, PCM core temperature, and PCM surface/fin-base temperature. The pyranometer was used to measure solar irradiance in the world, and this was recorded at the same time, which formed the foundation of computing the amount of solar energy received daily. The sensor readings were checked in the first phase of post processing against anomalies and outliers were identified as above 3 sigma. The sensor noise was reduced by a five minutes moving average filter without loss of transient features, including PCM melting plateaus. Hourly averages were subsequently derived to align with the interval of measurement of the distillate yield, and the thermal profiles could be directly correlated with the productivity of water. The productivity of distillate was measured in terms of hourly yield (mL/m^2/h) and cumulative yield per day (mL/m^2/day) divided by the basin surface area. The productive window was calculated by determining the time interval within which the yield of the distillate was greater than zero. This was the measure of the comparison between the extended operating hours offered by PCM and FPCM settings. The thermal efficiency was determined by the ratio of the useful latent heat in the collected distillate and the overall incident solar energy reaching the basin area.

2.4 Uncertainty Analysis

The uncertainties in experimental measurements are caused by the inability of the instruments to be precise, the calibration of sensors, and the variability of the environment. The uncertainty of each derived parameter was estimated by the root-sum-square (RSS) method, which is suggested by the ISO Guide to the Expression of Uncertainty in Measurement. The uncertainty U_R for a result R calculated from multiple measured quantities x_i is given by:

$$U_R = \sqrt{\sum \left(\frac{\partial R}{\partial x_i} \cdot U_{x_i} \right)^2} . \quad (1)$$

U_R is the total uncertainty of the result, R is the result calculated, x_i is each measured quantity and U_{xi} is the uncertainty of each measured variable.

For thermal efficiency (η), defined as:

$$\eta = \frac{m \cdot h_{\mathrm{fg}}}{I \cdot A \cdot t}, \quad (2)$$

where η represents the thermal efficiency, m is the mass of distillate (kg), h_{fg} represents the latent heat of vapourisation (kJ/kg), I is the solar irradiance (W/m^2), A is the basin area (m^2) and t represents the exposure time (s).

The following variables and uncertainties were considered. The key variables influencing thermal efficiency and their associated uncertainties are summarized in Tab. 2.

Table 2 Measurement variables, typical values, and associated fractional uncertainties

Variable	Typical value	Uncertainty (± fraction)	Description
Mass of distillate (m)	Variable (g)	±0.01 (1%)	From load cell or graduated jar (±1 g or ±1 mL)
Solar irradiance (I)	600–1000 W/m²	±0.03 (3%)	Pyranometer accuracy
Exposure time (t)	12 h	±0.005 (0.5%)	DAQ-controlled, manually verified
Latent heat of vaporization (h_{fg})	2260 kJ/kg	±0.002 (0.2%)	Slight variation due to temperature
Basin area (A)	1.0 m²	Negligible	Fabrication tolerance < ±0.001 m²

Therefore, the total uncertainty of thermal efficiency is of the order of ± 4.4% which is well within the acceptable limits of solar-thermal experiments carried out in the field.

In the daily productivity that is largely controlled by the quantified distillate volume and constant basin area of 1.0 m^2, the distillate volume is the main source of uncertainty. Using a measurement resolution of ±1mL or ±1g gives a resultant uncertainty of about 1%.

These results highlight the strength of the experimental rig and confirm the validity of the trends of thermal performance and freshwater yield observed in diverse solar-still systems.

3 RESULTS AND DISCUSSION

This section describes and discusses the experimental results of the three solar still configurations that were tested at different water depths. The main parameters of key performance, including productivity per day, thermal

efficiency, productive hours, and maximum hourly yield, are compared to assess the effects of PCM integration and fin enhancement.

3.1 Daily Productivity

Fig. 2 indicates the daily productivity of the three solar still configurations at various water depths. The findings indicate that there is a distinct difference between the traditional solar still (CSS), PCM-integrated still and FPCM-integrated still. CSS was the least productive in all the water depths with a yield of about 160-190 mL/m^2/day. This low output is explained by the fact that it does not store thermal energy and this results in extreme changes in temperature when there is a change in the amount of sunshine. This causes the CSS to produce water mostly on sunny days, and very little on a cloudy morning and evening. The addition of PCM gave some visible enhancement especially at the depth of 3 cm where the daily productivity was about 200-210 mL/m^2/day; this was a significant increase of about 15-20 percent relative to CSS. PCM served as a thermal reservoir, which stored unnecessary energy when the sun was blazing and then gave it back during the evening. Nevertheless, the PCM performance was not very consistent at 2 cm depth. Although the productivity was higher than with CSS, the difference was not as high as at the depth of 3 cm. This is attributed to the fact that PCM charging is not matched with the limited mass of water hence less latent heat is utilized. This has been also found in previous studies indicating that in cases of optimum PCM performance there is a tendency that water depth and storage capacity should be balanced.

The FPCM structure had the best productivity at all depths. At 2 cm deep, FPCM yielded approximately 325 mL/m^2/day, which is approximately 30-40 times better than CSS. Fins increased the thermal conductivity, which allowed heating water faster and the release of stored PCM heat more effectively. The concerted effort of PCM storage and fin-based conduction prolonged the distillation cycle and enhanced cumulative yield. Notably, although the experimental points are scattered, which is due to variations in irradiance per day and noise in measurements, the overall performance hierarchy (FPCM > PCM > CSS) did not change with depth.

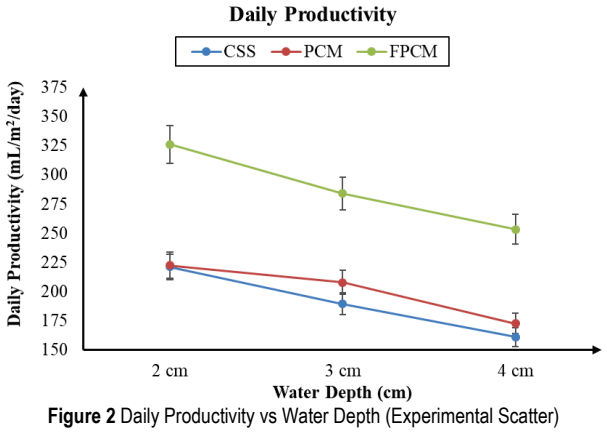


Figure 2 Daily Productivity vs Water Depth (Experimental Scatter)

3.2 Thermal Efficiency

One of the most significant indicators of solar still performance is thermal efficiency because it measures the proportion of incident solar energy that is converted to useful latent heat of vaporization. Fig. 3 shows how the thermal efficiency changes with the depth of water. Like the productivity trends, CSS had the least efficiencies of 1.62.3%. The resultant low efficiency is due to increased thermal losses due to the transparent cover and basin liner, and the lack of any storage mechanism. The PCM system was more thermal efficient with a range of approximately 1.8 - 2.2% with respect to depth. This is because PCM is capable of absorbing and storing latent heat, which then smooths out the thermal variations and keeps the water in the basin at relatively high temperatures even when it is not peak time. Nonetheless, the PCM performance at 2 cm depth was not as good as at 3 cm depth, as in the case of productivity. The discrepancy between PCM melting range and quick heating of shallow water constrained the contribution of PCM hence decreased the overall efficiency enhancement. This brings to the fore an important design consideration, which is that PCM properties should be chosen to match the operating water depths in order to be the most effective.

The FPCM set-up was always the most efficient, with 2.5 - 3.4%. Fins were also useful in enhancing thermal conductivity, decreasing the resistance between PCM and basin water. This enabled faster heat transfer when charging and more effective heat transfer when cooling resulting in increased evaporation rates. The dispersion of the efficiency values as shown by error bars is due to the real world uncertainties like the variability of the sun, the movement of the ambient air and sensor tolerances. However, the trend of the improvement of the efficiency was very consistent. The values are also similar to those that have been reported in earlier PCM and fin-enhanced solar still experiments thus strengthening the validity of the results.

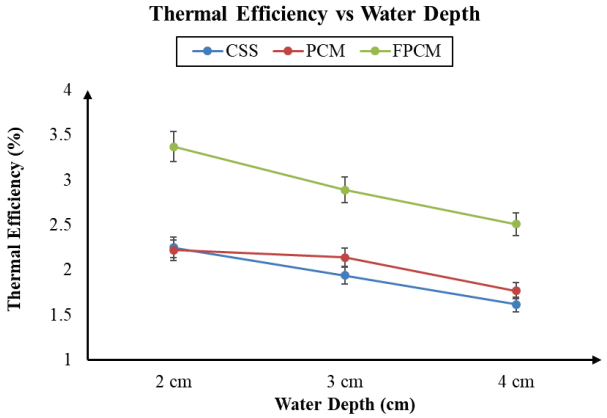


Figure 3 Thermal Efficiency vs Water Depth (Experimental Scatter)

3.3 Productive Hours

The productive hours or the time period during which the distillate is produced continuously gives an idea on the stability of the stills. As illustrated in Fig. 4, the CSS yielded

water under all depths over a period of approximately 11 hours, which is in close relation to the period of strong sunlight in the day. Once the sun went down, the CSS lost its thermal driving force very fast and production stopped nearly right away. The PCM-based setup prolonged the productive hours to 11-12 hours, especially at a depth of 3 cm. PCM accumulated heat at high radiation and slowly emitted it at night in the evening, continuing distillation after the sun went down. This sustained process brings out the buffering property of PCM. But at 2 cm depth the performance of PCM was not as effective. The productive hours of PCM at 2 cm were similar to CSS in a number of experiments. This anomaly has also been noted in earlier literature whereby there are thin layers of water, which heat excessively too fast to allow PCM to charge adequately and thus its contribution is limited. In this experiment, therefore, a balance between the mass of water and the storage capability was determined, and 3 cm depth was found to be more appropriate with PCM-based stills.

FPCM had the highest productive life, with the water output lasting almost 12 hours at all depths. The fins helped in effective PCM discharging, keeping basin water in evaporation favorable temperatures well into the evening. The error bars in Fig. 4 represent natural variability (±0.2 h) which is explained by variations in solar input and ambient conditions on a day-to-day basis. In general, the trend indicates that PCM and fins increase the working hours in contrast to CSS, and hence illustrate the feasibility of storage-enhanced designs.

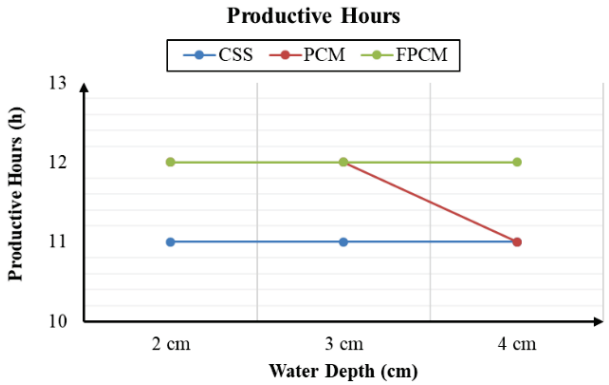


Figure 4 Productive Hours vs Water Depth (Experimental Scatter)

3.4 Peak Hourly Yield

Peak hourly distillate yield, which is shown in Fig. 5, indicates the immediate capacity of the stills to produce water at maximum radiation. CSS exhibited the lowest maximum yields with values of approximately 38 mL/h at depths of 2 cm decreasing with depth of water. PCM integration increased the peak yield marginally (approximately 32-35 mL/h) but its main benefit is the fact that it prolonged production time as opposed to maximizing the peak yield.

FPCM, in turn, showed a significant rise in peak yield, with the value of 50-55 mL/h at 2 cm depth. The fins increased the thermal conductivity, which made it easy to heat the basin water quickly and increase the evaporation when the sun radiated intensively. The depths of 3 cm and 4 cm reduced the maximum yield of FPCM as anticipated with increased mass of water and reduced heating speed, yet the relative excellence of FPCM was maintained.

Fig. 5 with its error bars (±5%) illustrates the variability of the experimental data because of the variations in solar radiation, wind effects, and errors of measurement of distillates by hand. Notably, the FPCM continued to achieve the highest peak yield at all depths, which validated the contribution of fins to the increase of system responsiveness. This observation is consistent with numerical predictions and previous experimental results in the literature, where fins always enhanced the rate of heating the basin and the maximum output capacity.

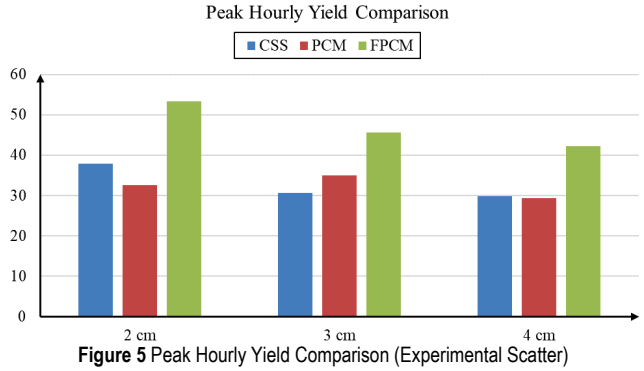


Figure 5 Peak Hourly Yield Comparison (Experimental Scatter)

The current experimental results are highly consistent with the results of the previous literature. As an illustration, Ahmed et al. [22] have recorded a 49.9% productivity enhancement due to finned PCM designs, and Surapararaju and Natarajan [15] have recorded a 24.26% productivity enhancement due to fin-integrated absorbers. Similarly, Sahu et al. [29] achieved a 46% improvement when using RT-58 PCM, and Mahala and Sharma [14] also reported similar efficiency improvements through PCM-based thermal storage systems. The finned PCM solar still in the present study demonstrated a relative 40 per cent increase in daily productivity and a 56% increase in thermal efficiency over the traditional still, which is within the range of improvement reported in the literature.

4 CONCLUSION

This paper has explored how the performance of solar still can be optimized by combining phase change materials (PCM) and heat transfer fins. Three configurations were done in both MATLAB-based numerical simulations and experimental trials, which are Conventional Solar Still (CSS), PCM-based Solar Still (PCM) and Finned PCM Solar Still (FPCM) with three basin water depths (2 cm, 3 cm, and 4 cm). It was aimed at measuring the improvement in the daily distillate productivity, thermal efficiency, productive hours, and peak hourly yield, and to compare the results of simulation and experimental validation.

The findings give the following conclusions:

The findings of the experiment showed that the CSS yielded the lowest yields of between 160 and 190 mL/m^2/day at depths. PCM integration improved productivity to 200 - 210 mL/m^2/day at 3 cm depth, which is equivalent to an increase of 15-20% compared to CSS. The FPCM was

always better than the two and had a depth of 2 cm, a rate of approximately 325 mL/m^2/day, or about 30-40 times higher than CSS. The relative performance hierarchy was similar to the experiment, although MATLAB simulations had a much higher absolute productivity of 7.3 L/m^2/day of FPCM at 2 cm depth. Experimentally determined thermal efficiencies were 1.6237 percent CSS, 1.822% PCM and 2.534% FPCM. The efficiency has been directly associated with the latent heat storage capacity of PCM and higher thermal conductivity of fins. Even though the absolute values were small in comparison with theoretical estimates, they agree with ranges in previous experimental results of solar stills.

CSS had a steady production of distillates of about 11 h/day, but only during direct sunshine. PCM was able to operate at 11-12 h/day, especially at 3 cm depth whereas FPCM had the longest productive window of around 12 h/day at all depths. The long running process is a direct confirmation of the importance of PCM in stabilizing the temperature of basin water and fins in facilitating the efficient discharge of heat. CSS was able to generate 38 mL/h of distillate at the maximum instantaneous rate, which was slightly higher than 3250mL/h in PCM and reached a maximum of 50-55 mL/h in FPCM at 2 cm depth. These benefits highlight the role played by fins in enhancing faster heating of the basin and increased evaporation during the peak sun hours.

This research contributes scientifically to the study of solar desalination by empirically showing how PCMs and finned structures can be used together to enhance the thermal efficiency and freshwater yield of solar stills. The results provide a practical model of maximizing the basin depth, thermal storage capacity, and heat-transfer configurations, and, consequently, improve the overall performance.

Socially, the outcomes support the creation of cost-effective and sustainable freshwater production technologies that can be used in rural areas and water-deficient regions. The research will increase access to clean water by enhancing the efficiency of solar stills with the help of low-cost materials, which will support the global sustainability and climate-adaptation goals.

5 REFERENCES

[1] Panagopoulos, A. (2021). Water-energy nexus: Desalination technologies and renewable energy sources. *Environmental Science and Pollution Research, 28*, 21009–21022. https://doi.org/10.1007/s11356-021-13332-8

[2] Esmaeilion, F., Ahmadi, A., Hoseinzadeh, S., Aliehyaei, M., Makkeh, S. A., & Astiaso Garcia, D. (2021). Renewable energy desalination; a sustainable approach for water scarcity in arid lands. *International Journal of Sustainable Engineering, 14*, 1916–1942. https://doi.org/10.1080/19397038.2021.1948143

[3] Sadaoui, Y., Belakroum, R., Cherrad, N., & Kadja, M. (2024). Experimental investigation of a novel hybrid solar-geothermal desalination process. *Desalination, 590*, 117982. https://doi.org/10.1016/j.desal.2024.117982

[4] Al-Addous, M., Bdour, M., Rabaiah, S., Boubakri, A., Schweimanns, N., Barbana, N., & Wellmann, J. (2024). Innovations in solar-powered desalination: A comprehensive review of sustainable solutions for water scarcity in the MENA region. *Water, 16*, 1877. https://doi.org/10.3390/w16131877

[5] Mandev, E., Muratçobanoğlu, B., Manay, E., & Şahin, B. (2024). Desalination performance evaluation of a solar still enhanced by thermoelectric modules. *Solar Energy, 268*, 112325. https://doi.org/10.1016/j.solener.2024.112325

[6] Kumar Pathak, P., Kumar, B., & Agrawal, R. (2025). Enhanced thermal performance analysis of single and double slope solar stills coated with TiO_2 nanoparticle-infused nanopaint. *Applied Engineering Letters, 10*, 146–159. https://doi.org/10.46793/aeletters.2025.10.3.3

[7] Fu, H., Dai, M., Song, H., Hou, X., Riaz, F., Li, S., Yang, K., Ali, I., Peng, C., & Sultan, M. (2021). Updates on evaporation and condensation methods for the performance improvement of solar stills. *Energies, 14*, 7050. https://doi.org/10.3390/en14217050

[8] Rajasekaran, A. K., & Murugavel Kulandaivelu, K. (2022). Performance comparison of solar still with inbuilt condenser and agitator over conventional solar still with energy and exergy analysis. *Environmental Science and Pollution Research, 29*, 83378–83388. https://doi.org/10.1007/s11356-022-21466-6

[9] Vigneswaran, S., Vijayarangan, B., Meikandan, M., Sakthivadivel, D., Kumaresan, G., Shanmugam, M., & Kumar, P. G. (2023). Economic and exergy analysis of passive solar still incorporated with an additional condensing surface material beneath the top glazing to enhance productivity. *Environmental Science and Pollution Research, 30*, 44861–44877. https://doi.org/10.1007/s11356-023-25495-7

[10] Rahmani, A., Khemmar, F., & Saadi, Z. (2021). Experimental investigation on the negative effect of the external condenser on the conventional solar still performance. *Desalination, 501*, 114914. https://doi.org/10.1016/j.desal.2020.114914

[11] Ganesan, K., Winston, D. P., Ravishankar, S., & Muthusamy, S. (2022). Investigational study on improving the yield from hybrid PV/T modified conventional solar still with enhanced evaporation and condensation technique—An experimental approach. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 44*, 5267–5286. https://doi.org/10.1080/15567036.2022.2083273

[12] Hassan, H., & Abo-Elfadl, S. (2021). Investigation experimentally the impact of condensation rate on solar still performance at different thermal energy storages. *Journal of Energy Storage, 34*, 102014. https://doi.org/10.1016/j.est.2020.102014

[13] Sathish Kumar, T. R., & Jegadheeswaran, S. (2022). Experimental investigation on finned solar still with enhanced thermal energy storage. *Journal of Thermal Science and Engineering Applications, 14*, 091001. https://doi.org/10.1115/1.4053228

[14] Mahala, T., & Sharma, N. (2024). Experimental investigations of a novel solar still with heat storage materials: Energy, exergy, economic and environmental analyses. *Desalination, 578*, 117467. https://doi.org/10.1016/j.desal.2024.117467

[15] Suraparaju, S. K., & Natarajan, S. K. (2021). Productivity enhancement of single-slope solar still with novel bottom finned absorber basin inserted in phase change material (PCM): Techno-economic and enviro-economic analysis. *Environmental Science and Pollution Research, 28*, 45985–46006. https://doi.org/10.1007/s11356-021-13495-4

[16] Essa, F. A., Mohammed, S. A., Basem, A., Alawee, W. H., Aljaghtham, M., Abdullah, A. S., Majdi, H. S., Sultan, A. J., Omara, Z. M., & Gamiel, Y. (2024). Maximizing desalination performance in pyramid distiller: Integration of vertical wick still and enhanced phase change material by nanoparticle and emerging fins. *Solar Energy, 283*, 113033. https://doi.org/10.1016/j.solener.2024.113033

[17] Agrawal, R., & Singh, K. D. P. (2021). Performance evaluation of double slope solar still augmented with binary eutectic phase change material and steel wool fibre. *Sustainable Energy Technologies and Assessments, 48*, 101597.

https://doi.org/10.1016/j.seta.2021.101597
[18] Kumar, A., & Kumar, R. (2024). Exergetic investigation and Taguchi-based optimization of a modified passive solar still augmented with nano-PCM & fins. *Journal of Energy Storage, 78*, 109935. https://doi.org/10.1016/j.est.2023.109935
[19] Omara, Z. M., Ahmed, M. M. Z., Alawee, W. H., Shanmugan, S., & Elashmawy, M. (2024). A comprehensive review of nano-enhanced phase change materials on solar stills with scientometric analysis. *Results in Engineering, 22*, 102088. https://doi.org/10.1016/j.rineng.2024.102088
[20] Kateshia, J., & Lakhera, V. (2022). A comparative study of various fatty acids as phase change material to enhance the freshwater productivity of solar still. *Journal of Energy Storage, 48*, 103947. https://doi.org/10.1016/j.est.2021.103947
[21] Khalilmoghadam, P., Kiyaee, S., Rajabi-Ghahnavieh, A., Warsinger, D. M., & Behshad Shafii, M. (2024). An improved passive solar still integrated with pulsating heat pipes and phase change materials. *Solar Energy, 275*, 112612. https://doi.org/10.1016/j.solener.2024.112612
[22] Ahmed, M. E. A. E., Abdo, S., Abdelrahman, M. A., & Gaheen, O. A. (2023). Finned-encapsulated PCM pyramid solar still – Experimental study with economic analysis. *Journal of Energy Storage, 73*, 108908. https://doi.org/10.1016/j.est.2023.108908
[23] Jahanpanah, M., Sadatinejad, S. J., Kasaeian, A., Jahangir, M. H., & Sarrafha, H. (2021). Experimental investigation of the effects of low-temperature phase change material on single-slope solar still. *Desalination, 499*, 114799. https://doi.org/10.1016/j.desal.2020.114799
[24] Aly, W. I. A., Tolba, M. A., & Abdelmagied, M. (2023). Experimental investigation and performance evaluation of an oval tubular solar still with phase change material. *Applied Thermal Engineering, 221*, 119628. https://doi.org/10.1016/j.applthermaleng.2022.119628
[25] Afolabi, L. O., Enweremadu, C. C., Kareem, M. W., Arogundade, A. I., Irshad, K., Islam, S., Oladosu, K. O., Elfaghi, A. M., & Didane, D. H. (2023). Experimental investigation of double slope solar still integrated with PCM nanoadditives microencapsulated thermal energy storage. *Desalination, 553*, 116477. https://doi.org/10.1016/j.desal.2023.116477
[26] Toosi, S. S. A., Goshayeshi, H. R., & Zeinali Heris, S. (2021). Experimental investigation of stepped solar still with phase change material and external condenser. *Journal of Energy Storage, 40*, 102681. https://doi.org/10.1016/j.est.2021.102681
[27] Dawood, M. M. K., Shehata, A. I., Shehata, A. S., Kabeel, A. E., Ramzy, K., Abdalla, A. M., Nabil, T., Elsabahy, M., & Elnaghi, B. E. (2022). Experimental investigation of a stepped solar still employing a phase change material, a conical tank, and a solar dish. *International Journal of Energy Research, 46*, 16762–16776. https://doi.org/10.1002/er.8337
[28] Hussen, H. M., Younes, M. M., Alawee, W. H., Abdullah, A. S., Mohammed, S. A., Atteya, T. E. M., Abbas, F., & Omara, Z. M. (2023). An experimental comparison study between four different designs of solar stills. *Case Studies in Thermal Engineering, 44*, 102841. https://doi.org/10.1016/j.csite.2023.102841
[29] Sahu, S. K., Natarajan, S. K., K, A. S., & Suraparaju, S. K. (2023). Experimental investigation of a solar still combined with phase change material (RT58) in southern India climatic conditions. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 23*7, 2635–2646. https://doi.org/10.1177/09544062221139986
[30] Agrawal, R., & Singh, K. D. P. (2022). Experimental investigation and computational modelling of double slope solar still integrated with eutectic phase change material. *Journal of Energy Storage, 52*, 104802. https://doi.org/10.1016/j.est.2022.104802
[31] Murali, G., Ramani, P., Murugan, M., Elumalai, P. V., Ranjan Goud, N. U., & Prabhakar, S. (2024). Improved solar still productivity using PCM and nano-PCM composites integrated energy storage. *Scientific Reports, 14*, 15609. https://doi.org/10.1038/s41598-024-65418-1
[32] Goshayeshi, H. R., Chaer, I., Yebiyo, M., & Öztop, H. F. (2022). Experimental investigation on semicircular, triangular and rectangular shaped absorber of solar still with nano-based PCM. *Journal of Thermal Analysis and Calorimetry, 147*, 3427–3439. https://doi.org/10.1007/s10973-021-10728-z
[33] Shalaby, S., Kabeel, A. E., Moharram, B. E., Shama, A., & Abosheiasha, H. A. (2022). Experimental study on the single basin solar still integrated with shell and spiral finned tube latent heat storage system enhanced by copper oxide nanoparticles. *Environmental Science and Pollution Research, 30*, 27458–27468. https://doi.org/10.1007/s11356-022-24104-3
[34] Elmghari, N., Rachidi, M. B., Salihi, M., Chebak, A., Al-Dahhan, M. H., & Chhiti, Y. (2025). Enhancing solar still productivity using phase change material: Experimental investigation under the climatic conditions of Benguerir, Morocco. *Applied Thermal Engineering, 269*, 126056. https://doi.org/10.1016/j.applthermaleng.2025.126056
[35] Boopalan, N., Kalidasan, B., Raj Kumar, D., Ragupathi, E., Gurumoorthy, M., & Praveenraj, S. (2021). Experimental study and analysis of single slope solar still integrated with phase change material. *IOP Conference Series: Materials Science and Engineering, 1059*, 012010. https://doi.org/10.1088/1757-899X/1059/1/012010
[36] Natarajan, S. K., Suraparaju, S. K., Elavarasan, R. M., Pugazhendhi, R., & Hossain, E. (2022). An experimental study on eco-friendly and cost-effective natural materials for productivity enhancement of single slope solar still. *Environmental Science and Pollution Research, 29*, 1917–1936. https://doi.org/10.1007/s11356-021-15764-8

Authors' contacts:

Kailas Deore, Research Scholar
(Corresponding author)
SSBT's College of Engineering & Technology,
Bambhori, Jalgaon, Maharashtra, India
kailasdeore1@gmail.com

Nilesh Salunke, Principal
SVKM's Institute of technology, Dhule, Maharashtra, India
nilesh.salunke@svkm.ac.in

Sudhakar Pawar, Professor
SSBT's College of Engineering & Technology,
Bambhori, Jalgaon, Maharashtra, India
Pawar.sudhakar@SSCOETjalgaon.ac.in

Yogeshkumar Pathak, Assistant Professor
R. C. Patel Institute of technology, Shirpur, Maharashtra, India
yogesh.pathak@rcpit.ac.in

Juber S. Khatik, Assistant Professor
R. C. Patel Institute of technology, Shirpur, Maharashtra, India
khatikjs@gmail.com

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20251118071140

Original scientific paper
Received: 2025-11-18, Accepted: 2026-01-05

Machine Learning-Based xG forecasting in the Top 5 European Soccer Leagues: A Comparative Analysis

Davronbek Malikov, Jaeho Kim*

Abstract: In the modern soccer world, also referred to as football, analytics has become integral IT serves a role akin to assistant coaches, contributing significantly to the analysis of team and individual player performance during both games and training sessions. Despite the abundance of advanced technologies, there is still a need for concrete measurements in soccer analytics. A prime example is Expected Goals (xG), a widely embraced metric that goes beyond mere scorelines to offer in-depth insights into player and team dynamics, and has proven its value in recent years. The identification of key features from a dataset to predict xG is a critical aspect of machine learning research applied to soccer analytics. However, prior investigations have either overlooked certain crucial features or failed to recognize their significance. This study proposes a novel approach by incorporating features related to coaches, coaching tenure, and tactics, which have the potential to improve predictive accuracy and provide actionable insights for teams and analysts. Using a dataset of 2,917 observations covering the top five European leagues, we applied regression-based machine learning models, employing a preprocessing pipeline and k-fold cross-validation to ensure robust evaluation. The findings reveal that the xG values for the English Premier League (EPL) surpass those of the other four leagues studied, with an average xG of approximately 1.93. This indicates that, on average, teams in the EPL tend to have a higher expected goal count per match than teams in other top European leagues. This superiority can be attributed to various factors, such as the higher average quality of teams and players, tactical nuances, and the overall competitiveness inherent within the EPL compared with the other leagues under scrutiny. In our study, we introduced new features such as coaches, coaching durations (years), and tactics. Incorporating these features enhanced the model performance across *MSE*, *EVS*, and R^2, thereby demonstrating the efficacy of our approach. Lasso and Ridge Regression models achieved improved predictive accuracy, with *EVC* and R^2 reaching up to 96%, while the Decision Tree model showed a nearly 6% reduction in the *MSE*.

Keywords: Expected Goals (xG); football; regression machine learning; Soccer; Top 5 league xG

1 INTRODUCTION

One notable difference setting soccer apart from other sports is the common occurrence of low-scoring games, where each goal is very important and can greatly affect the result of the match [2, 7, 23]. Consequently, the paramount goal of soccer is to successfully convert these opportunities into tangible goals, underscoring the importance of efficient goal conversion in the sport.

The Expected Goals (xG) metric quantifies the likelihood of scoring a goal by analyzing data from similar past shots. Opta is a renowned leader in sports analytics and provides comprehensive and detailed data for various sports [1]. Their extensive historical database, which contains nearly one million shots, offers valuable insights into player performance, team strategy, and match outcomes. By leveraging Opta's data, we can analyze and predict the probability of scoring opportunities, known as xG, with precision and accuracy. The xG values are assigned on a scale of zero to one. A value of zero indicates an impossible chance to score, whereas a value of one suggests that a goal is expected for every given chance [1]. For instance, a chance from the halfway line is less likely to result in a goal than that inside the penalty area. Using xG, numerical values can be assigned to these scenarios. If a chance from inside the box is given an xG of 0.1, this means that a player is expected to score one goal every ten shots in this situation. For example, if a player has 20 chances from inside the box in a single match, their expected goal total would be 20 times 0.1, or 2 xG [1]. The goal results of many matches, including those of the top five European leagues, were predicted. However, a high xG value does not always guarantee a soccer win. Consider a recent Premier League match between West Ham United and Brighton, as shown in Tab. 1. According to the xG statistics, West Ham had a probability of 0.83, whereas Brighton had a probability of 2.32. Logically, with an xG of 2.32, Brighton should have won the game. Surprisingly, however, the match ended in a 0:0 draw, meaning that no goals were scored despite high xG values.

Table 1 Example of xG and actual goals results (EPL match on 02.02.2024)

Team	xG	Actual goals
West Ham	0.83	0
Brighton	2.32	0

In this example, it is evident that the relationship between xG and actual goals is not always direct. Specifically, while a higher xG score indicates a greater likelihood of scoring for a given opportunity, it does not consistently translate to a higher number of goals. Similarly, an xG value of n does not necessarily imply scoring n goals in every instance. For example, consider a scenario in which a player has an xG value of 0.8 for a particular shot. This suggests that, on average, they have an 80 percent chance of scoring at that opportunity. However, in reality, they may fail to convert the chance due to factors such as exceptional goalkeeping, poor finishing, or a last-ditch defensive block. Conversely, a team may have a lower xG value for a match but still manage to score multiple goals. This could occur if they capitalize on fewer but higher-quality chances, or if they benefit from opposition mistakes that lead to goals. The accuracy of soccer xG predictions can be compromised by the selection and identification of key features. Furthermore, in soccer analytics, precise anticipation of goal-scoring opportunities is vital for strategic decision making and effective performance evaluation [15]. By integrating machine learning models for xG prediction, analysts and coaches can not only forecast potential scoring outcomes but also understand which features (player, tactical, and managerial) drive these predictions, providing a framework for explainable and data-driven decision-making in soccer.

However, previous research [17] has encountered limitations in accurately predicting xG owing to the exclusion of meaningful features that significantly influence xG outcomes. For instance, previous studies [17] have primarily focused on player position as the main feature.

However, team tactics, coaches, and coaching duration (years), which are closely related to player position, contribute to the xG prediction in our study. Including these features also enhances the interpretability of the model, as it reveals how managerial and tactical factors influence predicted scoring opportunities. These features are expected to improve xG prediction, because coaching decisions directly influence team strategy, player positioning, and shot selection. Therefore, the research gap lies in the limited incorporation of coaching-related features into the existing xG models. This study addresses this gap by integrating these features to improve prediction accuracy [17].

The primary objective of this study is to address this disparity by developing an advanced xG model that integrates novel and significant features. In light of the gaps identified in prior research, this study aims to develop a comprehensive framework for predicting xG by incorporating both traditional performance indicators and novel contextual features. Specifically, we hypothesize that the inclusion of variables such as coaching factors (e.g., coaching experience and duration) and tactical attributes, in addition to player- and match-level data, would significantly enhance the predictive accuracy of xG models. The overarching objective of this study is to demonstrate that a broader set of explanatory variables leads to more robust and reliable xG predictions, offering deeper insights into player and team performance evaluations. Additionally, our proposal involves comparing xG across the top five European football leagues: the English Premier League (EPL), German Bundesliga (Bundesliga), Spanish La Liga (La Liga), Italian Serie A (Serie A), and French Ligue 1. Such comparative analysis allows clubs and analysts to identify league-specific patterns, informing strategy and resource allocation across different competitive contexts. This new method aims to improve and customize the xG metric to the specific qualities of each league, thereby providing deeper insights into goal-scoring probabilities across various football settings. Our evaluation shows the predicted xGs for the top five leagues, and compares these values with the average actual goals for each league. We also identify the best xG players currently active in these top leagues. Overall, this study not only improves predictive accuracy for xG but also demonstrates how integrating novel features within ML models can support explainable and strategic decision-making in soccer.

The remainder of this paper is organized as follows: a review of related works, a section on data processing that outlines the steps taken to prepare the data for our study, an experiment section detailing our methodology, and a results section presenting the outcomes of our analysis by comparing recent results. Furthermore, the Top 5 xG section introduces a new category of xG designed for the Top Leagues, accompanied by a comparative evaluation. The paper concludes with a final section that summarizes our findings and insights.

2 BACKGROUND AND RELATED WORK

In this section, we present an outline of xG and its advantages for soccer teams, including coaches. In the following section, we integrate the relevant studies that encompass prior research.

2.1 Overview and Expected Goals

In recent years, football has undergone significant changes both on and off the pitch, with new tactics such as counterpressing and zonal marking altering its dynamics [13]. Alongside these tactical evolutions, advanced analytics has become increasingly prevalent, with xG emerging as a key metric for assessing player and team performance. xG provides a more nuanced understanding of how effectively teams convert their scoring opportunities and create high-quality opportunities. Moreover, the increasing use of data, particularly xG, has transformed how football matches are perceived and analyzed. While xG has become a central aspect of tactical discussions, it has also sparked controversy among fans. As mentioned earlier, xG is now an essential part of modern football analysis, influencing coaching decisions, player scouting, and even fan discussions on the game. For example, soccer team coaches consider xG metrics, and using xG statistics, coaches can prepare their teams according to the xG statistics for each player [6]. Let us look at the opinions of one of the coaches of the soccer team: "The xG of the two penalties is huge; one I didn't think was and the other was very soft," argued Newcastle manager Eddie Howe [12].

2.2 Related Work

Although the xG metric is gaining popularity in sports analytics [30], it is a relatively recent notation with a limited number of published papers [2, 26]. This can be partly attributed to the challenges associated with the collection of the experimental data. Privacy concerns make data collection a complex task, and the adoption of modern technologies for live data collection is not universally feasible for every soccer team because of the financial limitations they may face [16, 21]. Consequently, the widespread implementation of xG metrics poses challenges in sports analysis. While prior studies have explored xG computation using various machine learning and statistical methods, most focus predominantly on event and positional features, leaving managerial and tactical factors largely unexamined. This analytical comparison identifies a gap in feature selection that motivates our approach.

In soccer analytics, a closer examination of the historical progression of xG literature brings forth the notable work of the expected goals philosophy by James Tippett as a pioneering contribution [30]. Moreover, Tippett's research not only explains the development of the xG metric but also underscores its practical relevance to football. His study focused on the significant impact of xG on performance assessment and player recruitment (scouting), emphasizing its pivotal role in shaping strategic decisions within the complex domain of soccer analysis. Essentially, the philosophy of expected goals is a foundational piece that

explores the developmental journey of expected goals and their application within the intricacies of football.

Furthermore, in the computation of xG metrics, a combination of machine learning and statistical methods is required. Among these methods, logistic regression provides interpretable coefficients but may oversimplify complex interactions, whereas tree-based models and gradient boosting capture nonlinear relationships with higher predictive power. Neural networks offer flexibility for high-dimensional data but often lack interpretability. This comparison highlights that while predictive accuracy is important, understanding feature contributions remains a challenge. This collection of methodologies underscores the flexibility of choosing models for xG computation, showcasing the various paths available to enhance precision and efficacy in predictive analysis, including applications in player salary prediction [2, 20, 23, 31, 32, 36].

Moreover, the significance of shots as crucial elements in predicting xG metrics is emphasized. Several football studies reinforce the role of shots as effective proxies for success in sports [20, 34]. The conventional metrics of shots and shots on the target may not fully capture the intrinsic value of shot attempts. This research analyzes spatiotemporal patterns, specifically focusing on the 10-second build-up leading to shot attempts. The novelty of this study lies in introducing the concept of defender proximity to the shooter with the aim of enhancing the validity of shot assessments. Utilizing this approach, this study aims to evaluate shot quality and shed light on whether the success of a team in a game can be attributed to dominance or luck. Moreover, shooting type is an important feature for predicting the xG score [3, 32, 35] and distinguishing among four distinct shooting types: open-play footed shots, headers, free kicks, and penalty shots. This study constructed a multilayer perceptron to forecast shot outcomes using a dataset of approximately 10,000 shots [20]. Given its significance in elucidating the unpredictability of goal probability, shot location is a prevalent topic of discussion in almost all studies related to xG [19, 21, 27, 33, 34]. Previous studies primarily rely on event and positional features such as shot location, defender proximity, and type of shot [3, 20, 32, 35]. Few studies consider team-level contextual factors, and none explicitly incorporate managerial or tactical variables, indicating a gap that our study addresses.

Moreover, the techniques outlined earlier are not the only means of forecasting the outcomes using xG. For example, match–outcome prediction is an alternative approach in this context. In contrast to the anticipated goal outcomes, this subject has undergone a thorough examination in previous studies [8, 35]. The most frequently analyzed factor when forecasting match results is the concept of home advantage, which is a phenomenon observed in numerous sports. Home advantage was introduced as a categorical variable denoting home/away conditions. All three models consistently demonstrated that teams had a statistically higher likelihood of winning when playing at home than when playing away [18]. Moreover, an alternative approach employed randomness in match outcomes, with a substantial sample size of 7,304 matches spanning four seasons in the top five leagues. The variables considered in this analysis included distance, angle, rule setting, and body parts used in the shots [3]. An important factor to consider when examining match-outcome predictions is the current form of football clubs [14]. The study includes a variable defining the form of a team, representing the average result over the most recent n games, with 1 indicating a win, 0.5 for a draw, and 0 for a loss. Among the 24 form variables tested, evenly split between the home and away teams (12 each), 20 coefficients proved significant. Match outcome prediction studies focus on factors like home advantage, team form, or randomness in results [3, 8, 14, 18, 35]. Compared to xG-focused studies, these works provide insights into broader match-level outcomes but do not inform feature-level interpretability for shot-level probabilities. In contrast, our approach integrates tactical and coaching features to improve both prediction and explainability. These coefficients align with the expected sign direction–positive for the home team and negative for the away team–across varying values of n.

Generally, the features incorporated into these models are engineered from in-game data and divided into two main categories: event and positional. Event data includes detailed information about all pitch occurrences during a match, such as passes, duels, fouls, and shots. Each data point typically includes details, such as the location of the event on the pitch with x and y coordinates, the ending position of the event (for shots, passes, etc.), the player involved, the match in which the event occurred, the success or failure of the action, and various other variables. A team of individuals manually tags these data points while watching the game. In every machine learning project, the careful selection of features is a crucial aspect, and this also holds true for xG metric prediction. Feature selection can significantly affect the overall xG results. An illustrative study highlighted the importance of features such as goalkeeper positioning, player pressure radius, and opposition between the shot and goal in the development of xG [17]. Synthesizing these studies reveals a consistent gap: managerial and tactical features are largely omitted, limiting interpretability and actionable insights. Our study fills this gap by incorporating coaching experience, duration, and tactical attributes, providing a more comprehensive and explainable xG model.

Furthermore, our objective is to introduce the impact of incorporating new features into the development of xG metrics. Additionally, we present a new classification of xG metrics tailored specifically for the top five European leagues (e.g., La Liga xG).

3 HANDLING OUR DATASET

This section underscores the pivotal role that data play in our research efforts. We initiate the discussion by outlining the complexities of the data collection process, explaining both the origin of our data and pertinent details regarding specific features. The subsequent section expounds on the incorporation of supplementary data, thus enriching the comprehensiveness of our dataset exploration.

3.1 Data Sources

The significance of data in both machine-learning projects and research is evident. The quantity and quality of data play a crucial role in influencing the overall research

outcomes. The data used in this study were collected from two primary sources, FBref [9] and Statsbomb [29]. FBref is a comprehensive football statistics platform that provides detailed data on player and team performance, including metrics such as goals, assists, and passes. Statsbomb is another reputable provider of football data, offering advanced analytics and insights derived from detailed event data [29]. A comprehensive overview and additional details can be found on the publicly available website [28]. Moreover, we primarily utilized Statsbomb data in our dataset; however, to ensure consistency and accuracy, we also incorporated data from the FBref website. We gathered comprehensive and reliable data for our research by comparing and merging these two datasets. Furthermore, we augmented our dataset by incorporating data for new features acquired from a professional sports database, Transfermarkt [5]. In our study, we systematically collected and integrated data from Transfermarkt, focusing on the historical statistics and personal details of the players and coaches. This process involved aligning profiles using unique identifiers to ensure consistency in attributes, such as playing history, positions, team affiliations, and career paths. We chose Transfermarkt for its comprehensive performance data, which are essential for accurate xG forecasting, and for its detailed positional and league-specific insights, which help to account for variations in playing styles across Europe's top five leagues.

Table 2 Personal Details and Tactical Attributes of Coaches

Attribute	Value
Date of Birth/Age	Jun 10, 1959 (64)
Place of Birth	Reggiolo, Italy
Citizenship	Italy
Avg. term as coach	2.29 years
Coaching License	UEFA Pro License
Preferred formation	4-3-3 attacking
Agent	CAA Base Ltd. verified

This integration enhances our analysis by combining match-specific performance data with broader career statistics, providing a holistic approach to xG modeling. Tab. 2 presents new features derived from Transfermarkt data, exemplified by Carlo Ancelotti, Real Madrid's head coach for the 2021–2022 seasons. We enriched our dataset with coaching attributes, such as date of birth, nationality, UEFA coaching license level, average tenure, and preferred tactical formation. Incorporating coaching data is vital because the style, tactics, and experience of a coach significantly influence the xG of a player. By linking each player to their respective coach for a given season, we capture the impact of coaching strategies on individual player performance, providing a more comprehensive xG analysis. For instance, the average tenure of a coach indicates managerial stability, which influences player development and team roles. Preferred formations, such as a 4-3-3 attack setup, reveal the offensive or defensive strategy of a coach, which affects the goal-scoring opportunities of players. Linking players to their coaches allows xG forecasts to account for both tactical and managerial factors. By combining player and coach data, we enhance xG predictions by considering individual performance in broader tactical and managerial contexts.

3.2 Data Preprocessing

Our dataset comprises 2,917 observations, with each observation corresponding to a unique entity within the dataset. Tab. 3 provides an overview of the sample dataset, highlighting the key player attributes and coaching-related features. Despite the modest size of the dataset, 90 distinct features known for their reliability in predicting xG were carefully curated. Among these, xG serves as the primary target variable, representing the expected goals associated with each observation. The dataset includes player-specific attributes, such as name, nationality, position, age, birth year, affiliated squad, and competition level. The dataset primarily covers the 2020–2021 football season, with an average player age of 25.27 years. Feature selection follows domain knowledge and correlation-based criteria, while missing values are handled using mean imputation for numerical features and mode imputation for categorical features to maintain data completeness. A significant proportion of players come from top-tier European leagues, including the Bundesliga, Premier League, La Liga, Serie A, and Ligue 1. In addition, the dataset uniquely incorporates coaching-related factors, recognizing the influence of managerial strategies on xG outcomes. Features such as coach name, average tenure (Avg_duration_years), and preferred tactical approach (Tactic) provide insights into how different coaching styles impact players' xG. For example, players under possession-based attacking coaches may generate higher xG values, whereas those under defensive-oriented managers may limit offensive opportunities. These coaching and tactical variables are operationalized as contextual features that capture managerial influence beyond individual player statistics. Categorical attributes such as coach identity and tactical approach are encoded using label encoding to preserve model compatibility while maintaining relative distinctions across categories. Although label encoding does not impose ordinal meaning, it enables the models to learn systematic differences associated with specific coaching contexts. Average tenure is included as a continuous variable to reflect managerial stability and its potential effect on team cohesion and chance creation. Collectively, these features are intended to model structural and strategic environments that shape players' expected goal generation rather than to imply direct causal effects. To assess the predictive power of the model, machine learning metrics, such as R^2, were used to evaluate the performance across multiple models. Although the dataset provides valuable insights, future studies could benefit from a larger dataset to enhance model accuracy, improve generalizability, and capture deeper patterns within the data.

We systematically categorized the players into distinct positions: goalkeepers (GK), defenders (DF), midfielders (MF), and forwards (FW). Additionally, we identified players with adaptable capabilities, spanning positions such as Defender-Forward (DF-FW) and Forward-Midfielder (FW-MF). Moreover, coaching details, such as the average number of years a coach spends with one team, were considered crucial features in our dataset. The average coaching tenure within our dataset is 1.92 years, providing valuable insights into the stability and dynamics of coaching roles. Furthermore, it encompasses a diverse array of player

performance metrics within the realm of football, including, but not limited to, goal scoring, assist provision, defensive actions such as tackles and interceptions, passing accuracy, shooting endeavors, and involvement in penalty situations. In our data analysis process, we employed heatmap analysis, which is a visual technique that allows the identification of correlations between different features within the dataset. This method facilitates the exploration of complex relationships and aids in determining the importance of the features. By leveraging heatmap analysis, we sought to identify the most influential features, thereby enhancing the effectiveness of subsequent analyses and modelling efforts.

Table 3 Sample Dataset Featuring Player Information, Coaching Influence, and xG Metrics

Player	Pos	Age	Born	Comp	Coach	Avg_duration_years	Tactic	…	xG
Max Aarons	DF	21	2000	EPL	Dean Smith	2.44	Possession and attacking	…	0.7
Yunis Abdelhamid	DF	33	1987	Ligue 1	Oscar Garcia	0.79	Strong defense	…	1.2
Salis Samed	MF	21	2000	Ligue 1	Pascal Gastien	4.43	Strong defense	…	0.8
Laurent Abergel	MF	28	1993	Ligue 1	Christophe Pelissier	3.35	Possession and attacking	…	2
Charles Abi	FW	21	2000	Ligue 1	Pascal Dupraz	1.66	Strong defense	…	0

Table 4 Final Feature Overview of Features for the Dataset

Abbreviation	Description	Data Type	Note
Coach	Coach name	Object	New feature
Avg_duration_years	Avg_duration_years	Float	New feature
Tactic	Tactic description	Object	New feature
Succ	Success count	Integer	Based on correlation
At	Attempt count	Integer	Based on correlation
Cmp %	Completion percentage	Float	Based on correlation
Gls	Goals scored	Integer	Based on correlation
As	Assist made	Integer	Based on correlation
Pk	Penalty kick	Integer	Based on correlation
xG	Expected goals	Float	Based on correlation
SCA	Shot creation actions	Integer	Based on correlation
Field	Fielding count	Integer	Based on correlation
Def	Defensive actions	Integer	Based on correlation
GCA	Goal creation actions	Integer	Based on correlation
MP	Match played	Integer	Based on correlation
Min	Minutes played	Integer	Based on correlation
Crdy	Yellow cards	Integer	Based on correlation
Crdr	Red cards	Integer	Based on correlation
Sht	Shots taken	Integer	Based on correlation
Pass	Passes made	Integer	Based on correlation
Clr	Clearances	Integer	Based on correlation
Err	Errors committed	Integer	Based on correlation
Blocks	Blocks made	Integer	Based on correlation
Touches	Number of touches	Integer	Based on correlation

Tab. 4 presents the final set of 24 features selected through heatmap analysis with a correlation threshold of ±0.5, ensuring strong relationships with the target variable. This set includes three newly introduced features - coach, average coaching duration, and tactics - whereas the remaining 21 features were identified based on their correlation coefficients surpassing the predefined threshold. Each feature is detailed with its abbreviation, description, and data type, encompassing Object, Float, and Integer.

Categorical features, such as coach name and tactic description, are label-encoded into numerical values for model compatibility. Numerical features, including goals scored, assists made, and penalty kicks, are already in a suitable format for the model input. Integrating both categorical and numerical data provides comprehensive insights into player and team performance, thereby enhancing the predictive accuracy of the model. Moreover, to ensure the robustness of the encoded features, categorical variables such as coach and tactical approaches are transformed using label encoding, converting each category into a unique numerical value. We verified the number of unique categories for both the coaches and tactics.

3.3 New Features for Accurate xG

This study introduces three new features: Coach, Average Coaching Tenure, and Preferred Formation. These three features are considered significant and can assist in predicting xG. The soccer team coach plays a crucial role in shaping the playing style, strategy, and tactics of a team. Coaches have distinct philosophies and approaches to games that directly influence team performance on the field. By incorporating information about the current coach into the predictive model, we can capture the tactical preferences, game plans, and managerial decisions that affect the attacking dynamics of a team. For example, coaches play a pivotal role in shaping the identity and tactical approach of a team. Each coach brings forth a distinct blend of philosophy and strategy, molding the playing style of the team. Klopp and Simeone's contrasting approaches offer fascinating insights into the spectrum of coaching philosophies. Klopp's relentless pursuit of attacking football at Liverpool FC epitomizes his "gegenpressing" philosophy, in which high-intensity pressing and rapid counter-attacks reign supreme during the 2021–2022 season. Conversely, Simeone's

pragmatic approach to Atlético Madrid in the 2021–2022 season prioritizes defensive solidity and disciplined counterattack. Although Simeone's tactics may curb xG opportunities, his strategic prowess elevates Atlético's resilience and competitiveness on the field [25].

Moreover, the average coaching tenure refers to the typical duration in which coaches remain in charge of a team. Coaching stability and continuity have significant implications on team performance and player development. Teams with long-term coaches often benefit from consistent tactical frameworks, well-established play patterns, and strong cohesion. For example, Manchester United experienced coaching instability after the departure of Sir Alex Ferguson in 2013, with several managerial changes over the years [11]. In contrast, Real Madrid enjoyed a relatively stable coaching duration under Zinedine Zidane during the 2016–2018 seasons, which led the team to have multiple Champions League titles. The average coaching period reflects continuity and stability within the coaching setup of a team. Teams with consistent coaching leadership tend to develop a cohesive playing style, build long-term strategies, and foster strong team chemistry, all of which positively influence their xG predictions [22].

In addition, the formation adopted by a soccer team reflects its tactical setup and positional structure during matches. Different formations allocate players to specific roles and positions on the field, thereby affecting their interactions, movements, and attacking patterns. For example, FC Barcelona, particularly under Guardiola in the 2008–2012 seasons, often used 4-3-3 as its preferred starting shape, with an emphasis on intricate passing, positional interchange, and exploiting spaces between opposition lines. In contrast, Bayern Munich under Hansi Flick excelled with a more flexible and dynamic 4-2-3-1 or 4-3-3 formation in the 2019–2021 seasons, emphasizing quick transitions, direct attacking play, and aggressive wing play. The preferred formation shapes the tactical approach, player positioning, and offensive strategies of a team, directly influencing xG predictions. Understanding how teams adapt their formations to exploit their strengths and opponents' weaknesses enhances the accuracy of xG models by capturing the spatial dynamics and strategic variations inherent to different formations [4].

Finally, Tab. 5 shows the impact of coaches. Table 5 provides a comparative analysis of the statistical data, showing the differences in performance metrics for Liverpool before and after the tenure of Jürgen Klopp since 2015 [10]. Before Klopp's arrival, Liverpool played 350 games, winning 171, drawing 84, and losing 95 games. Their goals for and against were 586 and 384, respectively. This resulted in an average point per game of 1.7, win rate of 48.9%, and loss rate of 27.1%. The goals per game and goals against per game were 1.67 and 1.09, respectively. During this period, Liverpool won only one trophy–the League Cup.

However, under Klopp's guidance, Liverpool's performance transformed remarkably. In the same number of games (350), Liverpool won 210, drew 81, and lost 59 games. Their goals for surged to 727, and their goals against reduced significantly to 352. This yielded an impressive point per game average of 2.03, with the win rate increasing to 60% and the loss rate decreasing to 16.8%. Notably, the number of goals per game increased to 2.07, while goals against per game decreased to 1.005. Klopp's influence is further underscored by the significant increase in trophies won, with Liverpool clinching four prestigious titles: Premier League, Champions League, Super Cup, and FIFA Club World Cup. This stark contrast in Liverpool's performance before and during Klopp's tenure provides compelling evidence of the pivotal role played by coaches in shaping the performance and xG of a team. Klopp's tactical acumen, motivational prowess, and ability to instill a winning mentality not only elevated Liverpool's performance but also significantly impacted their ability to create and capitalize on goal-scoring opportunities. Thus, the data clearly illustrates how a coach can significantly influence team success, highlighting the importance of coaching in soccer.

Table 5 Comparison of Liverpool's Statistics before and with Klopp

Statistic	Before Klopp: until 2015	With Klopp: from 2015
Games	350	350
Won	171	210
Drawn	84	81
Lost	95	59
Goals for	586	727
Goals against	384	352
Points per game	1.7	2.03
Win rate	48.9%	60%
Loss rate	27.1%	16.8%
Goals per game	1.67	2.07
Goals against per game	1.09	1.005
Trophies won	1 (League Cup)	4 (Premier League, Champions League, Super Cup, FIFA Club World Cup)

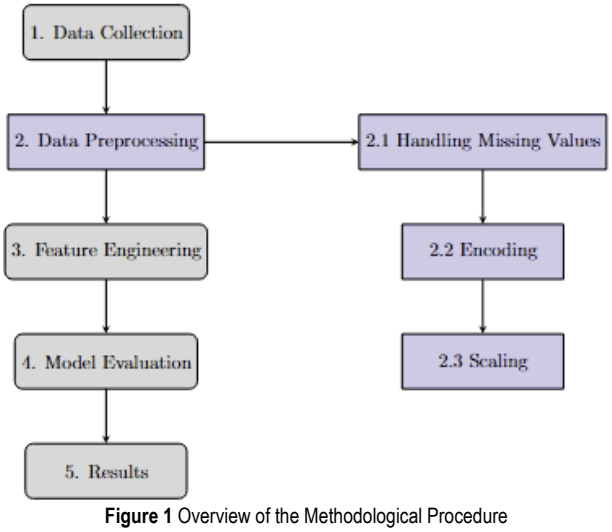


Figure 1 Overview of the Methodological Procedure

3.4 Overall Procedures for Our Analysis

This section describes the step-by-step process used in the analysis, as illustrated in Fig. 1. The methodology begins with data collection, in which Transfermarkt serves as the primary data source. After gathering the necessary data,

preprocessing is performed to ensure data quality and model readiness. Missing values are handled using the mean for numerical features and the mode for categorical features. Categorical variables are encoded using label encoding, whereas numerical features are subjected to standard scaling to maintain consistency across different feature ranges. Following preprocessing, feature engineering is conducted to extract meaningful information that enhances the predictive performance. Various machine learning models are then applied, as detailed in Section 4, in which the model comparison and evaluation are explained. Finally, the methodology concludes with the results presented in Section 5, which provides an in-depth analysis of model performance and findings.

4 MODEL COMPARISON AND EVALUATION

In this section, we commence by providing an introduction to various machine-learning models. Subsequently, we employed a comparative evaluation approach to select the most suitable model for our study, prioritizing the attainment of the highest score.

4.1 Model Introduction

In this study, we employed supervised learning methodologies because of the availability of both input and output (target) values. Within the realm of supervised learning, models typically fall into two categories: regression and classification. Notably, our study forgoes the utilization of classification machine learning models because our objective does not involve categorizing target variables or predicting binary outcomes, such as the presence or absence of a goal. Instead, we focused on predicting the xG score, which is a continuous numerical variable. Hence, we opted for regression machine-learning models to address this objective. These include the decision tree, which is valued for its interpretability and capability to capture complex relationships between features and target variables. The random forest model is well known for its ability to mitigate overfitting and enhance the predictive accuracy through ensemble learning. Ridge regression introduces a penalty term to address multicollinearity in the dataset and improve model stability. Lasso regression facilitates feature selection by penalizing the size of the model coefficients. Finally, the XGBoost regressor, a potent gradient boosting algorithm, is known for its ability to enhance predictive performance by iteratively improving weak models. The inclusion of managerial and tactical features is motivated by their influence on team strategy, player positioning, and shot selection. By integrating these features, the models aim not only to predict xG more accurately but also to provide interpretable insights into which factors drive expected goals, connecting the methodology to a decision-making framework.

4.2 Comparative Evaluation

Determining the most appropriate machine-learning model is a critical stage in experiments. Consequently, we employed a comparative evaluation approach to select the most suitable model for our dataset based on its highest score.

We utilized two experimental methodologies to compare the machine-learning models. In the first approach, we utilized a dataset obtained from Statsbomb [28, 29]. Conversely, in the second approach, we enrich our existing dataset by incorporating additional features such as coaches, tactics, and average duration. To achieve this, we sourced data from a professional soccer website called Transfermarkt [5]. Moreover, in both experiments, we evaluated the performance of our models using established metrics, including mean squared error (MSE), explained variance score, and R-squared. These metrics offer comprehensive insight into the accuracy, predictive capability, and overall fit of the models. The MSE quantifies the average squared difference between the predicted and actual values, providing a measure of the accuracy of the model. The explained variance score indicates the proportion of variance in the target variable that the model can explain, with higher scores reflecting a better predictive capability. R-squared measures the proportion of variance in the target variable explained by the model, and offers an indication of its overall fit to the data. Moreover, to ensure robust evaluation, we split the dataset into training and test sets using an 80/20 ratio, with the training set used to fit the models and the test set reserved for assessing the predictive accuracy of unseen data. To improve model performance, hyperparameters were selected using standard practices and preliminary tuning, including Randomized Search. Although the code cannot be shared due to restrictions, all model settings are fully described to ensure transparency and reproducibility. Performance metrics were calculated on the test set to provide a comprehensive assessment of predictive accuracy and model fit. Tab. 6 shows that the comparison between the performance of the regression models under two distinct approaches (without new features and with new features) reveals noteworthy improvements in the predictive accuracy and model performance with new features. Based on an evaluation of several regression models, the Ridge Regression model emerged as the best performer for our dataset when new features were incorporated. Here, we present a detailed analysis of the performance of each model and the rationale behind selecting the Ridge Regression model as the most suitable. The decision tree model shows improvement with the inclusion of new features, where MSE improves by approximately 6%, and explanatory power highlights minor changes in explanatory power and goodness of fit. This highlights the sensitivity of decision trees to additional contextual variables such as coach, coaching duration, and tactics. Ridge and Lasso Regression also show modest gains in fit, whereas Random Forest and XGBoost remain largely unaffected, with only negligible shifts in performance. Although the numerical improvements are modest, the results highlight how contextual variables such as coach, coaching duration, and tactics can shape predictive accuracy and model behavior. This contribution lies not only in performance gains but also in revealing which models are more sensitive to contextual information, offering valuable insights for both modeling and practical decision-making.

Table 6 Performance Comparison of Regression Models

Model	Without new features			With new features		
	MSE	*EVS*	R^2	*MSE*	*EVS*	R^2
Decision Tree	1.51	0.83	0.83	1.42	0.84	0.84
Random Forest	0.58	0.93	0.93	0.61	0.93	0.93
Ridge Regression	0.50	0.94	0.94	0.49	0.96	0.96
Lasso Regression	0.50	0.94	0.94	0.49	0.96	0.96
XGBoost Regressor	0.62	0.93	0.93	0.60	0.93	0.93

Overall, the inclusion of coaching- and tactics-related features improves the predictive performance across models, with the decision tree benefiting the most in terms of accuracy and explanatory capability. These results highlight the importance of feature engineering for enhancing model performance, particularly for models such as decision trees, which rely heavily on feature differentiation. Moreover, the adjustments implemented in approach two lead to all metrics being improved across various regression models compared to approach one. These improvements, reflected in the reduced *MSE* and increased *EVS* and *R*-squared scores, underscore the effectiveness of the approach in enhancing predictive accuracy and model performance.

5 EVALUATION RESULT

The results presented in this section are intended to provide a comparative and descriptive evaluation of predicted xG patterns across leagues and players. Our analysis focuses on identifying associations and performance trends rather than establishing causal relationships or statistical significance. In this section, we present an overview of our comparative evaluation of the top five soccer leagues based on their average xG using the Lasso Regression model, which is known for its superior performance metrics, as discussed earlier. Additionally, our analysis extends beyond a comparison of the top five leagues by incorporating individual-player statistics. The next section discusses the average xG values of the top five players. Tab. 7 shows the rankings of the top five football leagues based on their average xG during the 2020/2021 football season. Bundesliga, renowned for its fast-paced and competitive matches, leads the list with an average xG of 1.93.

Following closely, the Premier League is known for its attacking play style and high-scoring matches, with an average xG of 1.76. Ligue 1, characterized by its flair and technical prowess, maintains a solid position, with an average xG of 1.70. Serie A, renowned for its tactical sophistication and defensive solidity, ranks fourth, with an average xG of 1.65. Finally, La Liga, famous for its emphasis on possession-based football and technical proficiency, completes the top five with an average xG of 1.55. We present a comprehensive analysis comparing our predicted results with the actual goals scored across the top five football leagues during the 2022–2023 season [24]. The relationship between xG and actual goal-scoring outcomes is presented in Tab. 8. Moreover, comparing the predicted xG for the 2020/2021 season with the actual goals scored in the 2022/2023 season presents an opportunity to recognize the two-season gap. This temporal comparison is exploratory in nature and aims to examine the stability of xG-based patterns over time, rather than to claim longitudinal generalization or predictive causality across seasons.

Table 7 Comparison of Average Expected Goals across Leagues during the 2020–2021 Season

League	Average xG
Bundesliga	1.93
Premier League	1.76
Ligue 1	1.70
Serie A	1.65
La Liga	1.55

Table 8 Average Number of Actual Goals per Match in the 2022–2023 Season

Rank	Leagues	Actual goals
1	Bundesliga	3.17
2	Premier League	2.85
3	Ligue 1	2.80
4	Serie A	2.56
5	La Liga	2.51

For example, it offers the unique opportunity to observe the impact of a two-season gap. This comparative analysis allows for the long-term evaluation of the predictive model, assessing its consistency and reliability over time. By examining how well the model predictions hold across different seasons, researchers can gauge their robustness and applicability beyond their immediate performance.

Furthermore, it enables the assessment of performance consistency. Analyzing the predictions made two years prior to the actual outcomes helps determine whether the model consistently forecasts xG values effectively across various periods. Consistent accuracy reinforces confidence in the reliability of the model for decision making in soccer analytics. Finally, this approach validates the utility of the model in practical scenarios. Comparing predictions across seasons demonstrates how effectively the insights of the model translate into real-world applications, such as informing team strategies or evaluating player performance over an extended timeframe. Starting with the Bundesliga, our examination uncovers a notable average xG of 1.93 per match, highlighting a strong inclination to generate scoring opportunities within the league. This aligns seamlessly with the Bundesliga's renowned reputation for dynamic and attacking football, as evidenced by its leading average of 3.17 actual goals per match, demonstrating the efficiency of Bundesliga teams in capitalizing on these opportunities. When transitioning to the EPL, we observe a steady level of goal-scoring potential, with an average xG of 1.76 per match. This corresponds well with the robust average of the league of 2.85 actual goals per match, reaffirming its position as one of the most competitive and goal-rich leagues in the world. In Ligue 1, the typical xG hovers around 1.70 per game, indicating a moderate number of scoring chances. This aligns with the average of the league of 2.80 actual goals per match among participating teams, enhancing the sense of balance and excitement in French football. Moving to Serie A, we note an average xG of 1.65 per match, signaling a slightly reduced likelihood of goal-scoring opportunities compared with other leagues. Serie A's average of 2.56 actual goals per match indicates a marginally decreased goal-scoring rate relative to its xG expectancy. Lastly, La Liga exhibits the

lowest average xG per match at 1.55, suggesting a comparatively lower frequency of goal-scoring chances. However, the league maintains consistent goal-scoring efficiency, as evidenced by its average of 2.51 actual goals per match, showcasing effectiveness, despite fewer expected goal-scoring opportunities. These observed differences should be interpreted as correlational patterns rather than causal effects, reflecting league-specific playing styles and contextual factors captured by the model.

Table 9 Players and Average xG (in descending order)

Best xG Player	Average xG
Robert Lewandowski	32.6
Karim Benzema	23.5
Kylian Mbappe	23.4
Ciro Immobile	22.2
Mohammed Salah	21.8

Overall, our analysis underscores the robustness of our predictive model in capturing the dynamics of goal-scoring opportunities, thus providing invaluable insights into team performance and goal-scoring potential across top football leagues. Consequently, the results suggest that our model provides slight improvements in predicting xG relative to actual goal outcomes, offering insights into football dynamics at the highest level. Moreover, Table 9 presents the top five players ranked by average xG as estimated by our predictive model, illustrating how the model captures individual-level goal-scoring potential across the top five European leagues. These players consistently exhibit high predicted xG values, indicating that the model effectively identifies elite attacking performance patterns rather than relying solely on league-level trends. Robert Lewandowski records the highest average xG (32.6), reflecting the model's sensitivity to sustained shot quality and volume, and highlighting his central offensive role for Bayern Munich in the Bundesliga. Similarly, Karim Benzema (23.5) and Kylian Mbappé (23.4) demonstrate high xG values that align with their tactical roles and attacking responsibilities at Real Madrid and Paris Saint-Germain, respectively. Ciro Immobile (22.2) and Mohamed Salah (21.8) further illustrate the model's ability to generalize across different leagues and tactical systems, capturing consistent goal-scoring contributions in Serie A and the Premier League. By linking player-level xG outputs to the model's predictions, this analysis reinforces the internal coherence of the results and demonstrates how the proposed features support reliable xG estimation at both league and individual-player levels. While evaluating the model predictions across the top five European leagues, we observed some discrepancies between the predicted xG and actual goals. For instance, the Bundesliga shows the highest average actual goals per match (3.17), surpassing the Premier League (2.85) and La Liga (2.51). This outcome is influenced by the presence of a high-performing player, Robert Lewandowski, whose season average xG of 32.6 significantly elevates the overall scoring metric of the league. Similarly, although the Premier League and La Liga have strong offensive statistics, the predicted xG does not fully align with the actual goals in some matches, particularly when standout individual performance disproportionately affects league-level aggregates. These deviations indicate that the model may underpredict or overpredict matches with exceptional player contributions or atypical scoring patterns. Such an analysis highlights the fact that high-xG players or dominant performances in a season can skew league-wide statistics, emphasizing the importance of considering individual player influences in future model refinements.

6 LIMITATIONS AND FUTURE WORK

Although our study encountered certain limitations, particularly in data collection, constraints imposed by modern technology restricted us to specific online platforms, thereby reducing data diversity and potentially introducing biases inherent in the original collection methods. The dataset, which consists of nearly 3000 observations across five leagues, is relatively limited for machine learning models. However, representativeness is ensured by including all matches with complete player and event data during the study period, covering diverse teams, player roles, and match situations. The absence of real-time data hindered our ability to capture immediate trends and reduced the timeliness of the study. The temporal gap between the training (2020/21) and validation (2022/23) data may introduce confounding factors such as player transfers, managerial changes, and evolving team compositions. Accordingly, this cross-season comparison is interpreted as a robustness check rather than a controlled longitudinal or causal evaluation of the xG prediction model. The absence of real-time data and player injury information may have overlooked factors affecting the outcomes. Future research could leverage advanced data collection technologies, such as real-time tracking systems and integrate additional performance metrics to improve the predictive accuracy and applicability in football analytics.

7 CONCLUSION

In conclusion, this study introduces novel features to enhance the predictive capabilities of xG models in football analytics. The primary novelty of this work lies in the explicit integration of coaching-related and tactical variables into xG modeling, which are largely absent from existing event- and position-based approaches. Incorporating coach attributes such as average tenure and preferred tactics provides valuable insights into the multifaceted nature of xG prediction. Regression machine learning algorithms, including decision tree, random forest, ridge regression, lasso regression, and XGBoost regressor, have facilitated the development of robust predictive models. Comparative analysis indicates that the ridge and lasso regression models perform comparatively well, showing a lower MSE than other models, whereas the decision tree model demonstrates a 6% reduction in MSE with the inclusion of new features, highlighting the value of feature engineering and regularization techniques in enhancing prediction quality. The findings show that incorporating contextual and managerial factors into machine-learning frameworks can meaningfully extend conventional xG modeling approaches. Additionally, a statistical approach was proposed to compare the top five football leagues–Premier League, Bundesliga, La Liga, Serie A, and Ligue 1–based on xG metrics. The Bundesliga emerged as a standout league, exhibiting the

highest average xG values. The strong correlation between our predicted xG values and the actual goals per match in these leagues during the 2022–2023 season demonstrates the reliability and practical applicability of our model. This study also identified the top five players with the highest xG scores across these leagues, with Lewandowski leading. This comprehensive analysis offers valuable insights into league and player performance dynamics, and contributes to the advancement of football analytics and decision-making processes. The generalizability of the results is constrained by the dataset size, the exploratory nature of the cross-season comparison, and the absence of advanced model interpretability analyses, which are addressed as directions for future research. Overall, this research provides a reference point for informed decision making and supports further exploration in football analytics, acknowledging that model performance has certain limitations.

8 REFERENCES

[1] Analyst, O. (2023). Expected goals (xG). *Opta Analyst.* https://theanalyst.com/eu/2023/08/what-is-expected-goals-xg/

[2] Anzer, G., & Bauer, P. (2021). A goal scoring probability model for shots based on synchronized positional and event data in football (soccer). *Frontiers in Sports and Active Living*, 53. https://doi.org/10.3389/fspor.2021.624475

[3] Brechot, M., & Flepp, R. (2020). Dealing with randomness in match outcomes: How to rethink performance evaluation in European club football using expected goals. *Journal of Sports Economics*, 21(4), 335–362. https://doi.org/10.1177/1527002519897962

[4] Bundesliga. (2020). How Hansi Flick steered Bayern Munich to Klassiker victory over Borussia Dortmund in his first Bundesliga game. *Bundesliga.* https://www.bundesliga.com/en/bundesliga/news/how-hansi-flick-steered-bayern-to-4-0-win-over-dortmund-klassiker-tactics-8142

[5] Transfermarkt. (2023). Coaches: Transfermarkt data. *Transfermarkt.* https://www.transfermarkt.com/carlo-ancelotti/profil/trainer/523

[6] Driblab. (2020). Expected goals (xG): What it is and how it works. *Driblab.* https://www.driblab.com/driblab-en/expected-goals-xg-what-it-is-and-how-it-works/

[7] Eggels, H., Van Elk, R., & Pechenizkiy, M. (2016). Explaining soccer match outcomes with goal scoring opportunities predictive analytics. In *Proceedings of the Conference on Soccer Analytics.* https://doi.org/10.1080/000368400421101

[8] Falter, J.-M., & Perignon, C. (2000). Demand for football and intramatch winning probability: An essay on the glorious uncertainty of sports. *Applied Economics*, 32(13), 1757–1765.

[9] Sport Website. (2024). Soccer: FBref. *FBref.* https://fbref.com/en/

[10] Planet Football. (2022). Comparing Liverpool's record before and after Jürgen Klopp's appointment. *Planet Football.* https://www.planetfootball.com/quick-reads/comparing-liverpools-record-before-and-after-jurgen-klopps-appointment-in-2015/

[11] Mcevoy, S. (2018). Manchester United and Sir Alex Ferguson with 27 years of stability. *MailOnline.* https://www.dailymail.co.uk/sport/football/article-6507649/Man-United-27-years-stability-theyre-looking-FOURTH-boss-five-years.html

[12] Furniss, M. (2024). Liverpool 4–2 Newcastle stats. *Opta Analyst.* https://theanalyst.com/eu/2024/01/liverpool-4-2-newcastle-stats/

[13] Garratt-Stanley, F. (2022). Jobs in football: What is expected goals (xG)? *Jobs in Football.* https://jobsinfootball.com/blog/what-is-expected-goals-xg/

[14] Goddard, J. (2005). Regression models for forecasting goals and match results in association football. *International Journal of Forecasting*, 21(2), 331–340. https://doi.org/10.1016/j.ijforecast.2004.08.002

[15] Gu, C., De Silva, V., & Caine, M. (2024). A machine learning framework for quantifying in-game space-control efficiency in football. *Knowledge-Based Systems*, 283, 111123. https://doi.org/10.1016/j.knosys.2023.111123

[16] Herbinet, C. (2018). Predicting football results using machine learning techniques. *MEng thesis, Imperial College London.*

[17] Hewitt, J. H., & Karakus, O. (2023). A machine learning approach for player and position adjusted expected goals in football (soccer). *Franklin Open*, 4, 100034. https://doi.org/10.1016/j.fraope.2023.100034

[18] Joseph, A., Fenton, N. E., & Neil, M. (2006). Predicting football results using Bayesian nets and other machine learning techniques. *Knowledge-Based Systems*, 19(7), 544–553. https://doi.org/10.1016/j.knosys.2006.04.011

[19] Kharrat, T., McHale, I. G., & Pena, J. L. (2020). Plus-minus player ratings for soccer. *European Journal of Operational Research*, 283(2), 726–736. https://doi.org/10.1016/j.ejor.2019.11.026

[20] Lucey, P., Bialkowski, A., Monfort, M., Carr, P., & Matthews, I. (2015). Quality vs. quantity: Improved shot prediction in soccer using strategic features from spatiotemporal data. In *Proceedings of the Conference on Soccer Analytics.*

[21] Madrero Pardo, P. (2020). Creating a model for expected goals in football using qualitative player information.

[22] Marca. (2021). Zidane departs as Real Madrid's most successful coach. *Marca.* https://www.marca.com/en/football/real-madrid/2021/05/27/60af73fbca4741756c8b46da.html

[23] Noordman, R. (2019). Improving the estimation of outcome probabilities of football matches using in-game information. *Amsterdam School of Economics, Faculty of Economics and Business.*

[24] Okunev, L. (2023). Comparison of the top-5 football leagues by goals scored. *Medium.* https://medium.com/@okunevleo/%D1%81%D1%80%D0%B0%D0%B2%D0%BD%D0%B5%D0%BD%D0%B8%D0%B5-%D1%82%D0%BE%D0%BF-5-%D1%84%D1%83%D1%82%D0%B1%D0%BE%D0%BB%D1%8C%D0%BD%D1%8B%D1%85-%D0%BB%D0%B8%D0%B3-%D0%BF%D0%BE-%D0%B7%D0%B0%D0%B1%D0%B8%D1%82%D1%8B%D0%BC-%D0%B3%D0%BE%D0%BB%D0%B0%D0%BC-288e5ccba7a3

[25] Rajarshi. (2021). Jürgen Klopp vs. Diego Simeone: Head-to-head manager comparison. *Sportco.* https://www.sportco.io/article/jurgen-klopp-vs-diego-simeone-head-to-head-comparison-846574

[26] Robberechts, P., & Davis, J. (2020). How data availability affects the ability to learn good xG models. In *Machine Learning and Data Mining for Sports Analytics: 7th International Workshop (MLSA 2020), Co-located with ECML/PKD.* https://doi.org/10.1007/978-3-030-64912-8_2

[27] Schulze, E., et al. (2018). Effects of positional variables on shooting outcome in elite football. *Science and Medicine in Football*, 2(2), 93–100. https://doi.org/10.1080/24733938.2017.1383628

[28] Somdeep. (2023). Soccer dataset. *GitHub.* https://github.com/somdeep/Statball

[29] StatsBomb. (2023). StatsBomb: Soccer data. *StatsBomb.* https://statsbomb.com/what-we-do/soccer-data/
[30] Tippett, J. (2019). *The expected goals philosophy: A game-changing way of analyzing football.*
[31] Mead, J., O'Hare, A., & McMenemy, P. (2023). Expected goals in football: Improving model performance and demonstrating value. *PLoS ONE*, 18, e0282295. https://doi.org/10.1371/journal.pone.0282295
[32] Malikov, D., & Kim, J. (2024). Beyond xG: A dual prediction model for analyzing player performance through expected and actual goals in European soccer leagues. *Applied Sciences*, 14(22), 10390. https://doi.org/10.3390/app142210390
[33] Zeng, Z., & Pan, B. (2021). A machine learning model to predict player's positions based on performance. In *Proceedings of icSPORTS*, Valletta, Malta, 28–29 October 2021, 36–42. https://doi.org/10.5220/0010653300003059
[34] Anzer, G., & Bauer, P. (2021). A goal scoring probability model for shots based on synchronized positional and event data in football (soccer). *Frontiers in Sports and Active Living*, 3, 624475. https://doi.org/10.3389/fspor.2021.624475
[35] Malikov, D., Jung, P., & Kim, J. (2025). Predicting soccer player salaries with both traditional and automated machine learning approaches. *Applied Sciences*, 15(14), 8108. https://doi.org/10.3390/app15148108

Authors' contacts:

Davronbek Malikov
Department of AI Convergence Engineering, Gyeongsang National University,
501, Jinju-daero, Jinju-si, 52828 Gyeongsangnam-do, Republic of Korea
davronbekmalikov96@gmail.com

Jaeho Kim
(Corresponding author)
Department of AI Convergence Engineering, Gyeongsang National University,
501, Jinju-daero, Jinju-si, 52828 Gyeongsangnam-do, Republic of Korea
jaeho.kim@gnu.ac.kr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20251217001846

Original scientific paper
Received: 2025-12-17, Accepted: 2026-01-16

Speech Signal Enhancement Using Linear Predictive Coefficients Adaptive Filter

Rehab I. Ajel*, Tariq A. Hassan

Abstract: The linear predictive coding (LPC) is a technique that is widely utilized in speech processing, especially in speech spectral envelope modelling, where the aim is to estimate the vocal tract resonances. Consequently, the LPC is used in a vast range of applications, like speech coding, transformation, compression, and speech and speaker recognition. However, with all advances in speech signal processing techniques, the existence of background noise poses a compelling challenge and has a huge impact on the speech quality. This paper proposes an enhancement technique that benefits from the LPC coefficients and exploits them as parameters for an adaptive filter used in speech noise reduction. Unlike the other methods that used the LPC technique for vocal tract spectral envelope estimation or source residual modelling, the proposed method adopts the LPC coefficients to build a frame-dependent, bandpass-like adaptive digital filter that is completely derived from the LPC parameters. This filter will be generated from each speech segment and used to suppress additive noise while preserving real-speech components. The proposed technique is evaluated and compared with two traditional fixed-parameter filters, Wiener and Kalman filters, using speech samples of the TIMIT dataset. The samples are corrupted with additive Gaussian white noise with an input signal-to-noise ratio (SNR) of 20 dB. Practical results show that the performance of the proposed LPC-based adaptive filter in enhancing the speech signal is remarkable with the voiced speech segments, while producing comparable results to Wiener and Kalman filters with unvoiced speech. These points out the effect of frame-dependent LPC-based adaptive filter to deal with nonstationary signals like speech.

Keywords: adaptive filter; Linear Predictive Coding (LPC); low-pass filter; Kalman filter; speech enhancement; Wiener filter

1 INTRODUCTION

One of the most reliable and essential forms of human verbal message is speech. As well as, important information can be conveyed by the speech signal; such as linguistic content (the words being said, the language, and the grammar), emotional state, gender, and speaker identity [1]. One of the most popular frameworks adopted in the speech processing field is the source–filter model. In which, the speech is represented as a convolution of two major components: the source signal and the filter envelop.

The source signal approximates the oscillation in the vocal cords and dictates the fundamental frequency along with its harmonics. The filter envelope, on the other hand, describes vocal tract which composing the throat, mouth cavity, lips, and the nasal cavities. The spectral envelope of the speech signal represent the vocal tract shape or the filter envelope [2, 3].

Naturally, the speech signals are often vulnerable to various types of noise, such as channel noise, babble noise, and additive background noise that enormously affect speech clarity and quality.

During the production of the speech, a background noise is presented due to various conditions and states such as health and emotion, as well as due to low-quality communication channels or interferences with other signals [4].

As a result, speech signal enhancement becomes one of the most fundamental tasks in contemporary speech signal processing techniques.

Enhancing the audio signals by reducing or removing the unwanted noise in order to improve their clarity and quality is the major role of the audio enhancement techniques. The goal is to produce a clean signal for humans and ready for analysis by machines.

Noise reduction is crucial for many applications, such as speaker recognition [5], automatic speech recognition [6], emotion classification [7], and related speech-based systems.

In the past few years, many proposed approaches have been presented to enhance and improve the signal quality. Some adopt classical signal processing techniques, and others rely on data-driven methods.

For instance, deep denoising autoencoders have been used in learning the mapping between the noisy and the clean signals [8]. A wavelet-based method along with deep-learning for signal enhancement has been proposed in [9]. Swarm optimization, integrated particularly with convolutional neural networks, has been explored in [10].

Classic filtering approaches such as adaptive filters, Kalman filters, and Wiener filters are still excessively used because of their ability in noise reduction and signal enhancement [11–14].

Linear predictive coding (or LPC) has presented as one of the main techniques for speech enhancement. The LPC coefficients, as stated by the traditional LPC-based methods, are mainly adopted in spectral envelope estimation of the speech signal. Or, in other words, to model the vocal tract of the signal during speaking. Noise reduction, on the other hand, is basically accomplished via residual or source signal processing or via LPC-derived parameters embedding with Wiener or MMSE-based filtering frameworks. However, the LPC approach, in these methods, functions as an analytical procedure rather than serving directly in the denoising mechanism.

This paper proposes a different use of the LPC technique. As opposed to the traditional methods, the proposed work suggest to directly adopting LPC coefficients as frame-dependent adaptive filter weights for noise suppression and speech enhancement. In other words, rather than applying LPC exclusively for signal envelope approximation or residual or source signal modelling, the proposed method

builds an adaptive filter generated from the LPC coefficients, which will be continuously updated according to the noisy speech frame.

This technique will effectively contribute to signal noise removal without the need for residual reconstruction or the estimation of the noise power, as required in traditional Wiener or Kalman filtering methods.

The evaluation of our proposed system will be under the conditions of additive Gaussian white noise, and the practical testing results show that the LPC-based adaptive filter. The proposed system is evaluated under additive Gaussian white noise conditions, and experimental results demonstrate that the LPC filter-based approach exceeds (in some cases) the performance of the conventional Wiener and Kalman filters in terms of speech enhancement performance.

In this work, the main contribute can be outlined as follows:

1) A new adoption of linear predictive coding (LPC) coefficients as a frame-dependent weighted adaptive filter for explicit speech enhancement, instead of spectral envelope approximation or source signal modeling.
2) A time-domain LPC-based speech noise removal system signal transformation to reduce complexity and time consumption.
3) Evaluation of the proposed system performance demonstrates that the LPC–based adaptive filter can exceed the performance of other filters like Wiener and Kalman filters under different noises levels.

2 NOISE TYPES

Unwanted electrical or electromagnetic energy deteriorates data and signal quality. This is known as noise. Both digital and analog systems can experience noise, which can impact many file kinds and communications, such as text, programs, graphics, audio, and telemetry. The following are possible major impacts of noise on a given signal: information content, clarity, and loudness.

Generally speaking, three types of noise may be modelled that impact speech signals: impulsive, colored, and white noise. [15, 16]. A signal or sound that has all audible frequencies present at the same strength is known as white noise at every frequency. When two distinct noise signals are combined, the resulting noise signal has a power equal to the total of its power components. The broad-band spectrum of white noise gives it powerful masking powers.

Colored noise is any noise other than white noise. The frequency spectrum of colored noise is restricted within a range, in contrast to white noise. Various colored noises exist, such as brown, pink, orange, and so on, based on the noise's Power Spectral Density (PSD) gradation. Sudden noise bursts with a comparatively large amplitude are referred to as impulsive noise. In the signal of interest, this kind of noise results in click noises [17].

3 ADAPTIVE FILTER

In signal processing, the adaptive filter is a system in which the parameters are adjusted automatically in response to the input signal variation according to the surrounding environment. In contrast with the fixed-parameter filter (predefined filters), adaptive filters continuously upgrade their coefficients in accordance with an adaptation mechanism that is previously set in the system. The filter coefficients are updated to meet a selected performance criterion, which normally is an error signal.

Generally, the adaptive filtering technique is an algorithm that comprises the digital filter along with a refitting algorithm that manages the filter coefficient update [18].

Techniques like Least Mean Square (LMS), Normalized Least Mean Square (NLMS), and Recursive Least Squares (RLS) algorithms are normally used as adaptive algorithms. These algorithms are continuously update the filter coefficients until the optimal (or the best fit) solution is reached [19]. The main criterion of these algorithms is the errors between the desired version of the input signal (clean signal) and the filter output.

In this paper, a new strategy for adaptive filtering is proposed. In it, the LPC coefficients are adopted as parameters for the frame-dependent adaptive filter. This differs to the traditional error-driven algorithms normally used in this vicinity.

An adaptive all-pole filter with its parameters evaluated over time is constructed using the LPC coefficients of the short-term segment of the noisy speech signal. In this technique, the coefficients of the LPC-based filter are readjusted in accordance with the spectral characteristics of the speech signal that will help to estimate the enhanced version of the speech.

4 LPC PARAMETERS

Linear Predictive Coding (LPC) is a technique that is widely adopted in signal modelling and speech analysis, designed to efficiently approximate the spectral coefficients, or the envelope, of the speech signal.

One of the classical models of speech production is the source-filter model, where techniques like LPC, MFCC, and PLP are totally rely on this model. The speech, according to this model, is represented as a vocal tract all-pole filter output with a source signal excitation.

Because of its solid representation and approximate explanation, the LPC technique has been broadly adopted in areas like; speech analysis, coding, synthesis, and speech and speaker recognition [20].

Basically, any speech sample, according to the presumption of the linear prediction, can be represented as a mixture or combination of a pre-set number of the previous samples, by virtue of the contiguous speech samples high correlation. Mathematically, the speech sample $\hat{s}[n]$ can be approximated as:

$$\hat{s}[n] \approx \sum_{k=1}^{p} a_k s[n-k] \tag{1}$$

where p indicates the order of prediction and a_k represents the coefficients of the linear predictive.

The error between the predicted signal and the actual speech is called the residual or the excitation signal, mathematically is expressed as:

$$e[n] = s[n] - \hat{s}[n] = s[n] - \sum_{k=1}^{p} a_k s[n-k] \tag{2}$$

The aim is to determine the LPC coefficients $\{a_k\}$ by minimizing the mean squared error between the actual signal samples and the predicted ones. Normally, the autocorrelation-based technique, which employs the Levinson–Durbin algorithm, is used [21]. Modeling the resonances of the vocal tract requires such a stable all-pole filter established by this optimization.

The all-pole transfer function of the vocal tract can be expressed as:

$$H_{\mathrm{LPC}}(z) = \frac{1}{1 - \sum_{k=1}^{p} a_k z^{-1}} \tag{3}$$

According to this transfer function (or all-pole filter), the frequencies inherent with the speech signal will be concentrated while other spectral components (outside the speech envelope) will be attenuated. Therefore, the LPC-based adaptive filter will show a band-pass-like mode that keeps the most dominant speech frequencies.

In our proposed speech-enhancing system, the LPC-based filter coefficients will be determined based on a frame-by-frame scheme from the noisy signal and used as a parameters for a band-pass adaptive filter.

Each frame of the noisy signal $x[n]$ will pass through the LPC filter defined in (3) to estimate , to some extent, an enhanced (noise free) signal $\hat{s}[n]$.

For each frame, the LPC coefficients will continuously be modified, which will allow the filter to adapt in accordance with non-stationary spectral characteristics of speech to be able to suppress the additive noise without the need for residual or source signal reconstruction or stochastic noise modeling.

Assume that $x[n] = s[n] + v[n]$, denote a speech signal with additive white noise ($v[n]$). In our proposed framework, the speech signal should first be segmented into fixed unoverlabed frames. The LPC coefficients $\{a_k\}$ have to be estimated directly from each noisy frame. These coefficients are then utilized to generate an instantaneous adaptive band-pass filter. The LPC synthesis filter $H_{\mathrm{LPC}}(z)$ will be applied to the corresponding noisy speech frame in the time domain, all follow:

$$\hat{x}[n] = h_{\mathrm{LPC}} * x[n] \tag{4}$$

where h_{LPC} is the impulse response of the LPC filter and * is the convolution operation.

This filtering operation will selectively retain the speech formant- related frequency components while removing, to some extent, the noise-related components.

Fig. 1 shows one frame speech signal and its frequency envelope (formats) estimated using LPC coefficients.

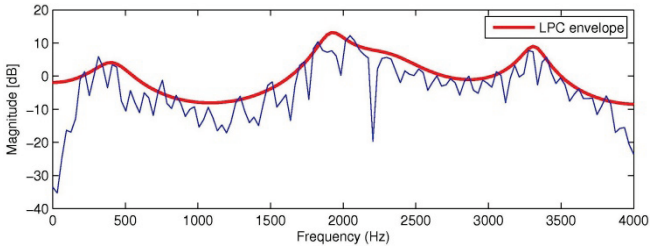


Figure 1 Speech signal frequency envelop and its formant

During the production of speech, the characteristics of the vocal tract resonances provoke a set of distinguished peaks in the envelope of the speech spectrum, which are known as the formant frequencies. The transferee function shows that the location of the resonance peaks is closely surrounded by the formant frequencies. The LPC with the all-pole filer framework can efficiently pattern this deed. The structure of the speech signal formant will be rounded up by the LPC filter poles.

This method takes advantage of the ability of the LPC to model the speech signal and acts as a band-pass adaptive filter that can modify its parameters according to time-varying speech characteristics in order to suppress the noise components that are located far from the modelled spectral structure.

5 THE PROPOSED MODEL

Our proposed system is to enhance and clarify the corrupted speech signal with additive white Gaussian noise. First, the input signal is segmented into short-time frames with a period of 20–30 ms. the speech characteristics tend to be quasi-stationary within this duration, so that the system can capture, almost, stable features of the speech.

To keep the frames in their original temporal structure, no windowing function is adopted.

Following segmentation, the next step is to classify the frames into speech-containing (or active) frames or silent frames. This classification is attained based on the calculation of frame energy. Practically, in noisy environments, the classification is highly challenging as the existence of background noise would make it hard to select the proper range of the energy threshold that adapts to separate the different frames. Yet, practically, an energy-based criterion was adopted to give an appropriate separation between these two types of frames. This would be appreciated for regulating the subsequent processing stages.

To eliminate the noise component within the high-frequency range, and prior to the LPC analysis step, each speech frame must pass through a low-pass filter. This is due to the fact that most of the speech energy is accumulated in a band within the low-frequency region of the speech, typically below 3.5 kHz. Additive noise, on the other hand, tends to

concentrate around high frequencies [22]. Because of its maximum flat passband and smooth frequency response, a sixth-order Butterworth low-pass filter with a cut-off frequency of 3 kHz is adopted in our system.

Following low-pass filtering, the next step is the LPC-based filtering scheme. For each frame of the noisy signal, the LPC procedure has to estimate the prediction coefficient set (order p=15 in our experiment) and use them to construct the LPC-based band-pass filter. These frame-dependent LPC coefficients are the main ingredients for the adaptive time-domain filters that serve to remove noise while maintaining the most speech-relevant components.

In this regard, we would like to draw attention to two important points: First, in the process, there is no overlap between the adjacent frames. In other words, the frames are processed sequentially to prevent discontinuity in the enhanced speech signal. Second, the continuous updating of the LPC coefficients with each frame will enable the adaptive filter to adapt to the non-stationary features of the speech signal.

The general processing steps of the proposed LPC-based speech enhancement model is illustrated in Fig. 2.

6 EXPERIMENTS AND RESULTS

As explained in the earlier sections, the input speech (noisy signal) will first be segmented into fixed-length frames (in our experiment, between 20 and 30 ms duration) as a preprocessing stage of our proposed speech enhancement scheme. The short-time strategy setup will enable the system to take advantage of the semi-stationary nature of speech within a short period of time while allowing frame-dependent adaptation of the LPC-based filter.

Before starting with LPC analysis, the high frequencies of the input frames should be attenuated using a low-pass filter, which will assist in noise reduction as the additive noise normally concentrates within the high-frequency area of the speech. Butterworth low-pass filter is used due to its steady frequency response and maximally flat passband. Based on this step, the system will seek the superiority of speech-relevant parameters, including frequency envelope or formant structure, within the low-frequency part of the spectrum.

The LPC-based adaptive filtering approach is now ready to extract the clean and clear signal out of the noisy frames. The LPC- based filter will be examined with different levels of additive white noise. The main objective of this empirical evaluation is to investigate the extent to which the LPC parameters can effectively participate in constructing a highly potent adaptive filter. This would help to eliminate the undesirable additive noise while retaining the speech characteristics.

Important notes have to be mentioned about the speech signals used in our system: First, all signals have the same sampling frequency, to ensure that the system stays consistent for all inputs. Second, the frames are evaluated separately according to whether they are voiced or unvoiced since they have different spectral characteristics and need a bit different management for noise removal. Also, the practical analyses focus on examining the ability of the proposed system to secure the real speech component while suppressing the noise artifacts.

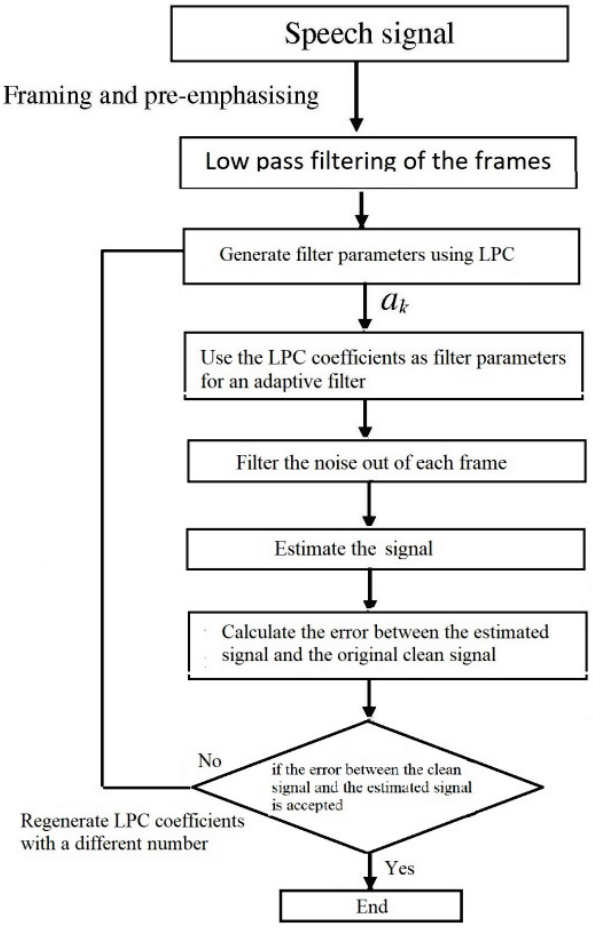


Figure 2 Block diagram of the suggested filtering model

This section shows the performance evaluation of the proposed framework for voice and unvoiced speech segments with a different range of the additive Gaussian white noise, presenting the level of enhancement in speech quality under noise effect.

The setup started with Fig. 3, which shows depictive examples of both voiced and unvoiced speech frames in clean and noisy states in the time domain. The frames are taken from the BT Millar speech database of the spoken words "one" and "six". Both are corrupted by additive Gaussian white noise at a ratio of 50%. The figure also emphasizes the explicit temporal typical feature of voiced and unvoiced speech segments and the noise impact on each type.

The corresponding frequency-domain depiction for these frames is presented in Fig. 4 illustrate the spectral distinction between voiced and unvoiced speech as well as the noise additive effect on the speech spectrum.

Fig. 5 illustrates the effect of low-pass filtering (at 3 kHz cut-off) on the noisy speech spectrum. It is clear how the low-pass filter significantly attenuates the unwanted noise.

The LPC scheme is applied to each low-pass filtered frame to estimate the speech spectral envelope. The extracted envelope is consequently utilized to construct a band-pass-

like filter. Fig. 6 shows the voiced and unvoiced spectra together with their corresponding LPC envelopes. It is very clear how the spectral envelope is concentrated around the resonance or the formant frequency of the speech.

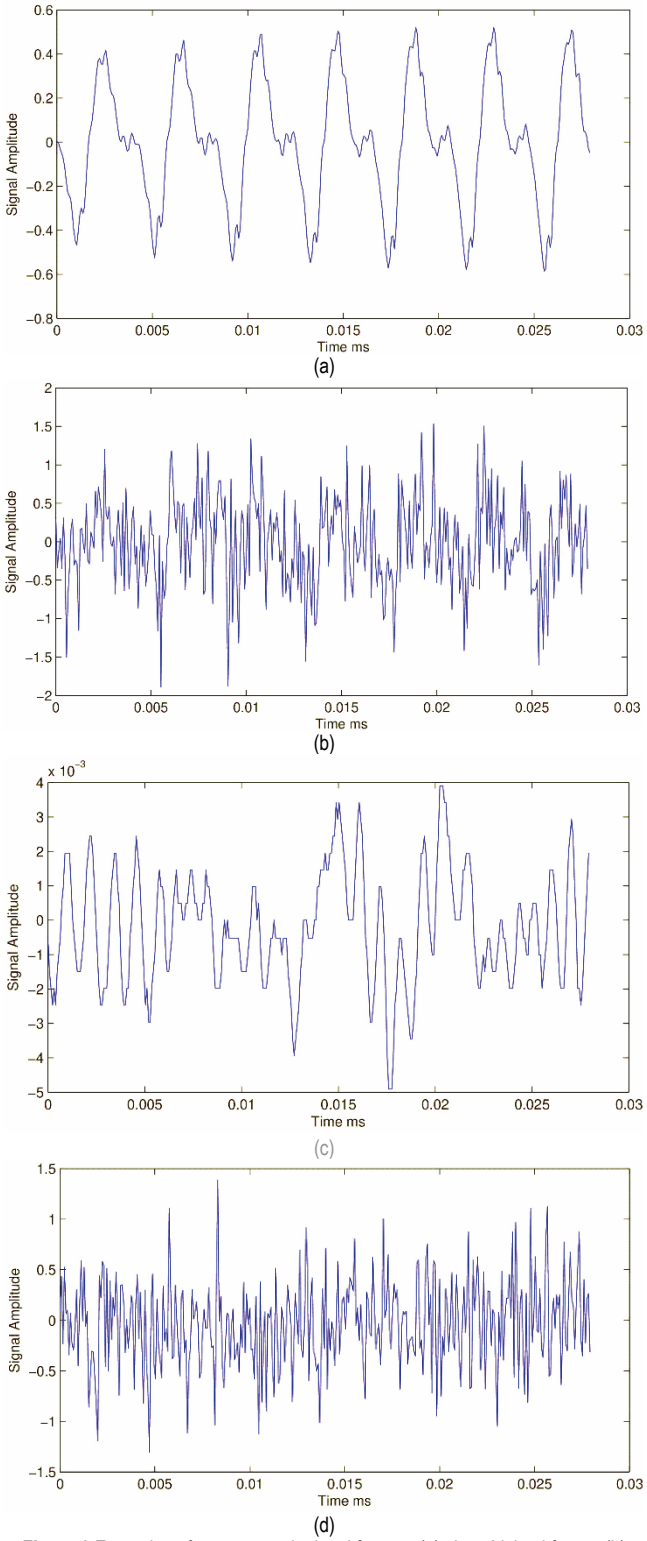


Figure 3 Examples of some speech signal frames: (a) clean Voiced frame, (b) Voiced frame with additive noise, (c) clean unvoiced frame, (b) Noisy unvoiced frame.

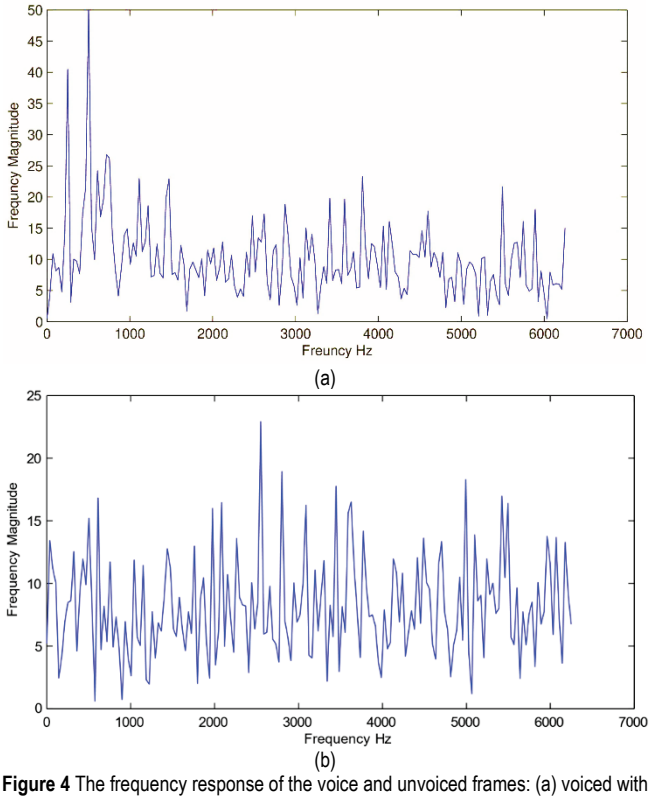


Figure 4 The frequency response of the voice and unvoiced frames: (a) voiced with additive noise frame, (b) unvoiced frame with noise.

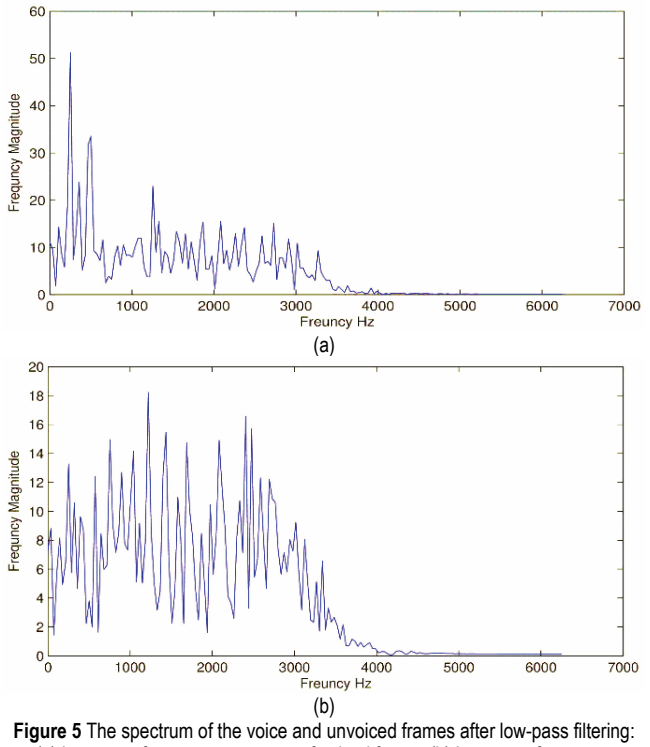


Figure 5 The spectrum of the voice and unvoiced frames after low-pass filtering: (a) Low-pass frequency response of voiced frame, (b) Low-pass frequency response of an unvoiced frame.

As mentioned earlier, the key concept for the proposed framework is to take advantage of the LPC coefficients to construct a band-pass filter that is able to adapt its parameters according to the frame in hand. The idea is that the bands

have to concentrate around the formant frequencies and to suppress unwanted spectral components related to the noise.

The potential strength of the proposed method lies in the way of processing each frame. It utilizes a frame-depending strategy in extracting the real-speech part of the corrupted signal. This proposed strategy enables the system to generate a new adaptive filter for each speech frame, which in turn permits the system to look distinctly at each corrupted frame in order to generate the proper filter.

This is particularly crucial for signals (like speech), whose spectral features vary rapidly over time.

To evaluate the performance of the proposed system, standard objective and perceptual metrics, such as Signal-to-Noise Ratio (*SNR*) and Peak Signal-to-Noise Ratio (*PSNR*) (expressed in decibels (dB)), and Mean Squared Error (*MSE*), are used to qualify the system's capability to enhance the speech signal buried under additive noise.

Mathematically, these equations can be expressed as:

$$MSE = \frac{1}{N}\sum_{n=1}^{N}\left(s[n]-\hat{s}[n]\right)^2 \tag{5}$$

$$SNR(\text{dB}) = 10\log_{10}\left(\frac{\sum s^2[n]}{\sum\left(s[n]-\hat{s}[n]\right)^2}\right) \tag{6}$$

$$PSNR(\text{dB}) = 10\log_{10}\left(\frac{(\max s[n])^2}{MSE}\right) \tag{7}$$

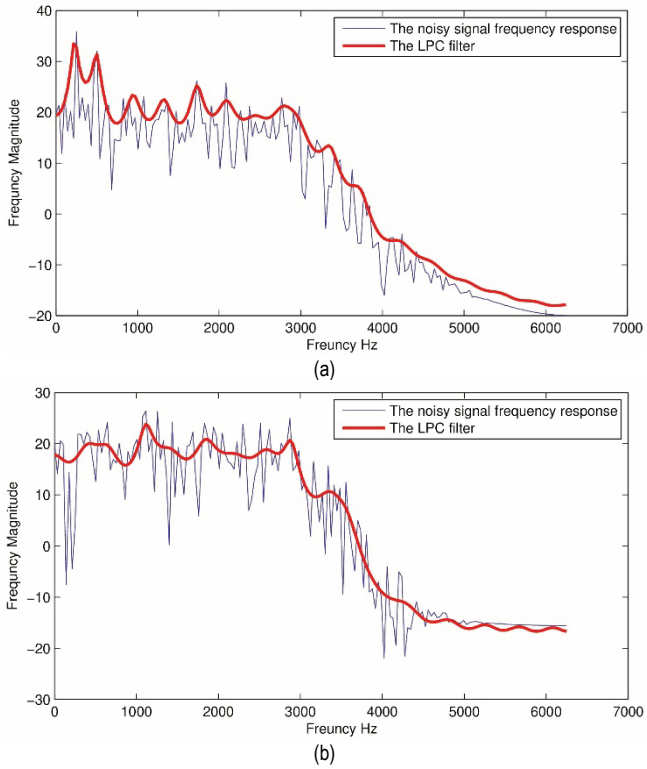


Figure 6 The frequency response and LPC envelope for voiced and unvoiced frames in db scale: (a) The voiced frame, (b) The unvoiced frame.

Fig. 7 presents a voiced and unvoiced corrupted frames with additive Gaussian white noise along with their corresponding enhanced frames achieved using the LPC-based adaptive filtering procedure with prediction order (p = 25).

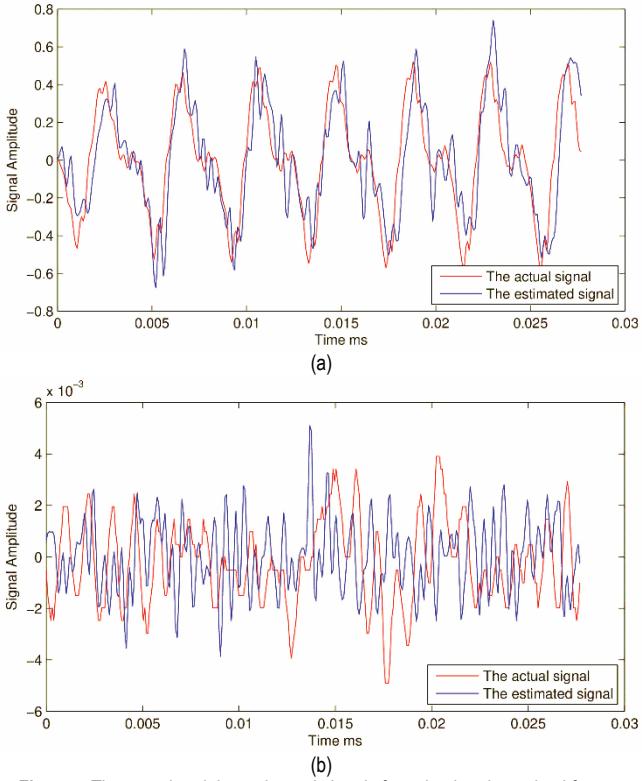


Figure 7 The actual and the estimated signals for voiced and unvoiced frames utilizing the LPC-based adaptive filter: (a) The voiced frame, (b) The unvoiced frame.

The presented results are for frames with 25 ms duration and Gaussian white noise with an input *SNR* of 20 dB.

As shown in the figure, using the proposed system, the obtained *PSNR* is approximately 85% for the voiced speech and about 50% for the unvoiced speech. This indicates that the frames with voiced tones can be handled well by the LPC-based filter compared to the unvoiced speech.

Fig. 8(a)–(d) summarizes the results of speech estimation and enhancing the voiced and unvoiced frames using Wiener and Kalman filtering approaches. The assessment is done under the same noise conditions (*SNR* = 20 dB).

Using the Wiener filter, the obtained PSNR values are approximately 75% for the voiced speech and 40% for the unvoiced speech. Equivalently, the Kalman filter achieved *PSNR* values of approximately 77% for voiced speech and 42% for unvoiced speech.

According to the obtained results, we can note that all filters can enhance the corrupted signal, but with different levels of accuracy.

The problem is that the Wiener and Kalman filters are relying on fixed filter coefficients (or pre-set parameters) that make them unable to cope with the signal's unpredictability and unstable changes with time. So the need for an instantaneous-adaptation filter that is able to modify its

parameters according to the signal at hand is very essential for unstable and highly fluctuated signal like speech.

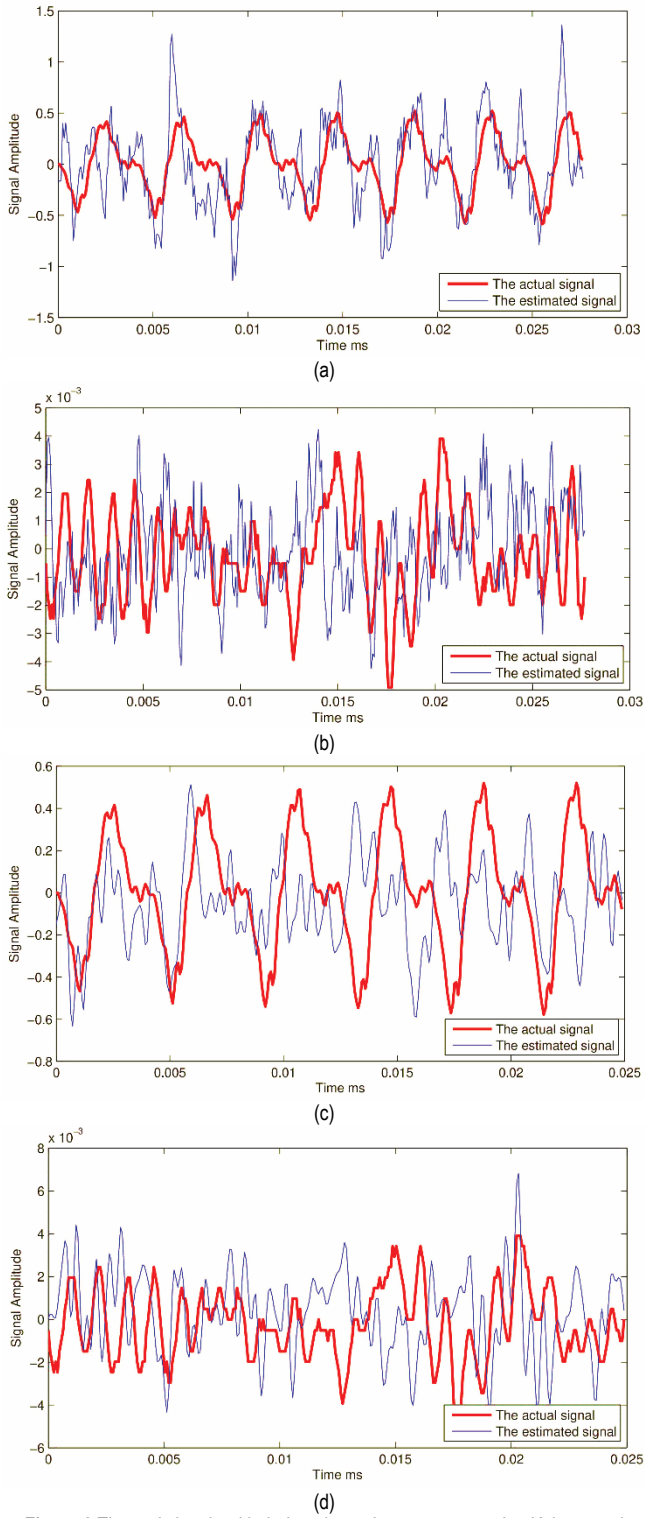


Figure 8 The real signals with their estimated counterparts using Kalman and Wiener filters: (a) real and estimated signal for voiced speech using Kalman filter, (b) real and estimated signal for unvoiced speech with Kalman filter, (c) real and estimated signal for voiced speech using Wiener filter, (d) real and estimated signal for unvoiced speech using Wiener filter.

Quantitatively, in the case of voiced speech, the mean squared error (*MSE*) values are approximately 0.0092, 0.018, and 0.024 for the LPC-based adaptive filter, Kalman filter, and Wiener filter, respectively. This demonstrates the exceptional performance of the proposed LPC-based approach in retaining voiced speech characteristics compared with other fixed-parameter filters.

As indicated in Figs. 7 and 8, the three filtering techniques, LPC-based, Kalman, and Wiener, have the ability to suppress the noise and enhance the speech signal with different levels of accuracy. The status is that nearly all methods exhibit excellent performance in terms of voiced speech. However, with the unvoiced speech frames, they all show limited performance.

Actually, we can note that the performance decline of the LPC-based technique to enhance the unvoiced speech attribute to the fact that this system is almost totally relying on the formant frequency envelope or the structure of the vocal tract. Since, it's well-know, that the unvoiced speech lacks a well-defined formant structure it's nearly noise, like speech.

Consequently, the LPC-based filter, particularly in this case, will be unable to distinguish between the real speech and the additive noise. Thus, the performance will decline dramatically.

We can attribute the remarkable performance of the LPC-based adaptive filter, especially with the voiced speech, to many key factors. First, it is different from the traditional Wiener and Kalman filters that totally utilize a previously set, unchangeable, optimized parameters.

The proposed method, on the other hand, builds up a dedicated filter that adapts its parameters depending on the frame it processes. This frame-wise adaptation allows the system to better track and identify the local spectral characteristics of the speech signal.

Second, the coefficients of the adaptive filter are estimated directly from the signal under analysis. This will allow the system to take advantage of the intrinsic acoustic information carried by the speech segment. Thus, the produced filter would be highly correlated with the instantaneous spectral envelope, or vocal tract resonant, of the speech signal, particularly in the case of voiced frames.

Nevertheless, the progress in the system performance is less noticeable in the case of unvoiced speech, which is because of the absence of the formant structures that adversely affect the LPC-based formants spectral modelling.

7 CONCLUSION

The paper presents three strategies for speech enhancement and background noise suppression: the LPC-based adaptive filter, the Wiener filter, and the Kalman filter. The latter filters are famous and widely used in the signal processing field, especially in signal enhancement and noise reduction. However, these filters are performed well under stationary noise conditions since they utilized a fixed or pre-set filter parameters that will never be changed or updated during the signal process. This will reduce their ability to tune with the nonstationary nature of speech signals.

The experimental results show that a frame-dependent adaptive filter with parameters that are directly derived from the speech signal provides excellent enhancement performance particularly for the voiced segments of speech. In other words, the proposed LPC-based adaptive filter technique exceeds the traditional fixed-parameter filtering approaches in noise suppression while retaining real-speech structure.

These practical results emphasize the importance of a signal-dependent adaptive filter for noise reduction and speech enhancement. Also, they point out the significant characteristics of the speech that can be carried by voiced segments. Consequently, LPC-based adaptive filtering could be a promising substitute to conventional fixed-parameters filters (Wiener and Kalman filters) for enhancing nonstationary signals.

8 SYSTEM LIMITATIONS AND PLANS FOR FUTURE

Although the proposed system produces imperfect performance with unvoiced segments of speech, practically, due to the absence of formant structure that is normally connected with the voiced speech. This will restrict the efficacy of the LPC-based filtering. Future work needs to focus on producing hybrid approaches that take advantage of the LPC-based filtering and prominent features of the unvoiced speech.

Acknowledgment

My colleagues and I in this article would like to thank the University of Mustansiriyah for supporting this work and the reviewers for providing useful suggestions, allowing for the improved presentation of this work.

9 REFERENCES

[1] Haque, M., & Bhattacharyya, K. (2018). Speech background noise removal using different linear filtering techniques. In S. Bhattacharyya, T. Gandhi, K. Sharma & P. Dutta (ed.), *Advanced computational and communication paradigms* (pp. 297–307). Springer Singapore. https://doi.org/10.1007/978-981-10-8240-5_33

[2] Dutoit, T. (1997). Linear prediction synthesis. In *Speech synthesis* (pp. 201–228). Springer Netherlands. https://doi.org/10.1007/978-94-011-5730-8_8

[3] Migimatsu, K., & Tokuda, I. (2019). Experimental study on nonlinear source–filter interaction using synthetic vocal fold models. *The Journal of the Acoustical Society of America, 146*(2), 983–997. https://doi.org/10.1121/1.5120618

[4] Proakis, J. G., & Manolakis, D. G. (1996). *Digital signal processing: Principles, algorithms, and applications* (3rd Ed.). Prentice Hall.

[5] Taherian, H., Wang, Z.-Q., Chang, J., & Wang, D. (2020). Robust speaker recognition based on single-channel and multi-channel speech enhancement. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 28*, 1293–1302. https://doi.org/10.1109/TASLP.2020.2986896

[6] Ochiai, T., Iwamoto, K., Delcroix, M., Ikeshita, R., Sato, H., Araki, S., & Katagiri, S. (2024). Rethinking processing distortions: Disentangling the impact of speech enhancement errors on speech recognition performance. *arXiv*. https://arxiv.org/abs/2404.14860 https://doi.org/10.1109/TASLP.2024.3426924

[7] Wijayasingha, L., & Stankovic, J. A. (2021). Robustness to noise for speech emotion classification using CNNs and attention mechanisms. *Smart Health, 19*, 100165. https://doi.org/10.1016/j.smhl.2020.100165

[8] Lu, X., Tsao, Y., Matsuda, S., & Hori, C. (2013). Speech enhancement based on deep denoising autoencoder. In *Proceedings of Interspeech 2013* (pp. 436–440). https://doi.org/10.21437/Interspeech.2013-130

[9] Gutiérrez-Muñoz, M., & Coto-Jiménez, M. (2022). An experimental study on speech enhancement based on a combination of wavelets and deep learning. *Computation, 10*(6). https://doi.org/10.3390/computation10060102

[10] Karthik, A. M. I. J. (2023). An optimized convolutional neural network for speech enhancement. *International Journal of Speech Technology, 26*(4). https://doi.org/10.1007/s10772-023-10073-6

[11] Haque, M., & Bhattacharyya, K. (2018). Speech background noise removal using different linear filtering techniques. In S. Bhattacharyya, T. Gandhi, K. Sharma & P. Dutta (Ed.), *Advanced computational and communication paradigms* (pp. 297–307). Springer Singapore. https://doi.org/10.1007/978-981-10-8240-5_33

[12] Das, O., Goswami, B., & Ghosh, R. (2016). Application of the tuned Kalman filter in speech enhancement. In *IEEE First International Conference on Control, Measurement and Instrumentation (CMI2016)* (str. 62–66). https://doi.org/10.1109/CMI.2016.7413711

[13] El-Fattah, M., Dessouky, M., Abbas, A., Diab, S., El-Rabaie, E.-S., Al-Nuaimy, W., Alshebeili, S., & Abd El-Samie, F. (2014). Speech enhancement with an adaptive Wiener filter. *International Journal of Speech Technology, 17*, 53–64. https://doi.org/10.1007/s10772-013-9205-5

[14] Sailaja, B. L. V., & Sunitha, P. (2019). Wiener filtering in frequency domain to enhance speech corrupted by colored noise. *International Journal of Recent Technology and Engineering*.

[15] Kondoz, A. M. (2004). Efficient LPC quantization methods. U *Speech coding and synthesis* (Chapter 5, pp. 87–147). John Wiley & Sons. https://doi.org/10.1002/0470870109.ch5

[16] Gold, B., & Morgan, N. (2011). *Speech and audio signal processing: Processing and perception of speech and music*. John Wiley & Sons. https://doi.org/10.1002/9781118142882

[17] Oudre, L. (2015). Automatic detection and removal of impulsive noise in audio signals. *Image Processing On Line, 5*, 267–281. https://doi.org/10.5201/ipol.2015.64

[18] Haykin, S. S. (2014). *Adaptive filter theory*. Pearson.

[19] Kalamani, M., & Krishnamoorthi, M. (2021). Modified least mean square adaptive filter for speech enhancement. In N. Dey (Ed.). *Applied speech processing* (pp. 47–73). Academic Press. https://doi.org/10.1016/B978-0-12-823898-1.00004-7

[20] Makhoul, J. (1975). Linear prediction: A tutorial review. *Proceedings of the IEEE, 63*(4), 561–580. https://doi.org/10.1109/PROC.1975.9792

[21] Huang, X., Acero, A., & Hon, H.-W. (2001). *Spoken language processing: A guide to theory, algorithm, and system development*. Prentice Hall PTR.

[22] Das, D. (2022). A systematic review of the effects of noise characteristics on human mental performance. In D. Chakrabarti, S. Karmakar & U. R. Salve (Ed.), *Ergonomics for design and innovation* (pp. 1705–1717). Springer International Publishing. https://doi.org/10.1007/978-3-030-94277-9_145

Authors' contacts:

Rehab I. Ajel
(Corresponding author)
Department of Computer Science, College of Science,
Mustansiriyah University,
Palastine St, 10064 Baghdad, Iraq
E-mail: rihabajel@uomustansiriyah.edu.iq

Tariq A. Hassan
Department of Computer Science, College of Education,
Mustansiriyah University,
Palastine St, 10064 Baghdad, Iraq
E-mail: tariq-mono2016@uomustansiriyah.edu.iq

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240703184715

Preliminary communication
Received: 2024-07-03, Accepted: 2025-07-03

Utilising Data Centres' Side Streams – A Waste Heat Case Study

Juuso Kivijakola*, Silja Heikkinen, Pekka Tervonen, Harri Haapasalo

Abstract: This article investigates the different challenges and phases of building an ecosystem to utilise data centre waste heat. Data centres consume an increasing amount of energy, and currently, in many cases, excess heat is not utilised beyond heating the data centre property. The research examines how a circular economy can be integrated into the operations of a data centre's waste heat business ecosystem and how a business ecosystem based on utilising waste heat from a data centre should be built. The research is based on a literature review, interviews with experts and stakeholders in the case ecosystem and a survey. Based on the research, a concept for a business ecosystem based on the utilisation of waste heat from a pre-existing data centre perspective is modelled, and the concept outlines things to be considered when building such an ecosystem. Key findings were the challenges, opportunities and prerequisites of developing a data centre waste heat ecosystem and the roadmap of constructing one.

Keywords: business ecosystem; circular economy; data centre; side streams; value chains

1 INTRODUCTION

The world's economy still relies heavily on a linear business model lacking built-in mechanisms for reuse or recycling, which is not sustainable in the long run because of environmental degradation and resource depletion [1]. A linear economy is based on material consumption, where the resources are used to make products, and after use, they become waste [2]. The endless use of limited resources is affecting nature, and the circular economy is thought to be one solution to these environmental challenges [3, 4]. The circular economy is a model in which as few resources as possible are used, and these resources are then reused to prolong their life cycle. At the same time, waste generation decreases [5]. The purpose of the circular economy is to avoid creating waste by using it as a resource in other processes [6].

The EU directive [7] states that data centres with a combined rated power above 1 MW are required to repurpose their waste heat for heating or other energy recovery uses, except when doing so is not technically or economically viable. This approach supports a circular economy and helps decrease reliance on fossil fuels. Yet currently, it is common for data centres to not utilise the waste heat in any way [8]. The amount of data and internet usage is increasing rapidly; thus, the number of new data centres worldwide is growing [9, 10]. Furthermore, data centres are vast consumers of electricity, estimated to be up to 1.5% of the global energy demand [11]. Because of this, the environmental effect of the data centres' operations is significant and should be considered [10].

This study aims to create a concept for utilising waste heat from a data centre in a business ecosystem. Typically, 40% of a data centre's energy consumption is attributed to cooling [10]. This means that the northern parts of Europe and other countries with colder climates are propitious locations for building data centres because of the reduced amount of cooling required. Data centres also create business opportunities and employment in sparsely populated areas, which in the case of Nordic countries have a lot of renewable energy and a reliable infrastructure required by the data centres. The research goals were as follows: To determine where data centres' waste heat could be utilised and what affects the creation of the business ecosystem, as well as to publish the information for operators.

These goals were condensed into the following research questions:

RQ1. What are the challenges and opportunities of utilising data centre waste heat?

RQ2. What are the main steps in creating a business ecosystem for the utilisation of data centre waste heat?

In addressing these questions, a comprehensive understanding was obtained regarding the factors involved in integrating sustainable development, particularly the circular economy, into business practices. Additionally, insights were gained on how to utilise waste heat generated by data centres and how these concepts are realised within a business ecosystem for waste heat. The concept was developed from the perspective of an existing data centre, indicating that it cannot be directly applied in all aspects when constructing a new data centre waste heat business ecosystem, but it includes considerations that should be taken into account.

2 LITERATURE REVIEW

2.1 Side Streams

The European Parliament and Council's Waste Framework Directive [12] defines how to protect the environment and people's well-being by preventing waste and transitioning to a circular economy. The directive presents a five-step waste hierarchy that defines the order of priority for actions to prevent the generation of waste based on waste management legislation. The five steps are waste prevention, preparation for reuse, recycling, recovery and disposal. In nature, no waste is produced, but everything serves as a resource for other processes [3]. For that reason, in a value chain based on a circular economy, side streams arising from processes are recovered and utilised to increase value according to the waste hierarchy [13].

A large amount of potentially reusable waste and side streams remain unutilised. Industrial side streams include everything that results from production processes, except for

the main product [14]. By further processing and combining side streams of different industries, new products can be created [15] and new sources of income can be found [14]. A company's sustainable development efforts should be integrated into its value chain through a life cycle approach to products, which includes increasing the use of recycled materials in product design, reducing energy consumption in production and properly managing hazardous waste after the materials' useful life [16].

Industrial symbiosis refers to a group of companies (for example, clusters or supply chains) where the sharing of resources promotes the efficiency of several stakeholders, either by reducing, for example, energy use and costs, or by offering waste (side streams) from one company's process as a contribution to the value creation process of another company [13]. In industrial symbiosis, companies utilise each other's waste and resources in their operations [2]. Generally, the development of products manufactured through industrial symbiosis encourages material recycling and reduces resource consumption and waste generation [15].

2.2 Data Centres

Because of the digitised society, data centres have become a critical part of the infrastructure [17]. To meet the growing needs of the internet and cloud computing, the number and size of data centres have increased in recent years [9, 10]. Data centres are constantly growing larger, becoming more complex and using a lot of electricity due to the increased demand for data storage and computing power [18].

Data centre equipment is generally divided into IT equipment and support equipment [17]. IT devices include servers, storage devices and communication networks, which directly affect the main functions of the data centre, such as data storage, processing and transferring. Support equipment usually consists of the cooling system and electrical infrastructure as well as lighting and equipment related to safety and other support functions [17].

The substantial energy consumption and associated costs of data centres primarily stem from the electrical demands of servers and cooling systems in addition to lighting, power distribution and other auxiliary operations [19–22]. Many data centre companies also consume excessive energy to ensure high reliability for their clients, resulting in low utilisation rates of IT systems [23]. In other words, although servers and storage devices have become increasingly energy-efficient during active use, they operate at full capacity for only a limited time [17].

2.3 Energy Efficiency in Data Centres

A data centre is an environment where precisely defined temperatures and cooling needs for information technology, such as servers and computers, are controlled and maintained [24]. Nearly all the electricity used in a data centre ultimately becomes heat, which must be extracted to maintain proper operation [17]. Maintaining a suitable temperature is important, as IT equipment can be damaged or stop working at too high a temperature or due to rapid temperature fluctuations [9].

Data centres consume a significant amount of energy, which leads to CO_2 emissions, and the energy consumption of servers and computers in data centres is constantly increasing [21]. Consequently, energy consumption and environmental impacts have developed into significant challenges in the design and operation of data centres [19], particularly the high energy consumption of cooling systems [10]. As current equipment increasingly generates heat loss, the efficiency and management of energy use play a crucial role in system design [21]. Additionally, indoor thermal conditions significantly impact the electricity consumption of air-conditioning systems [9].

According to the IEA [11], data centres account for approximately 1%–1.5% of global electricity use and, together with data transmission networks, 0.6% of global greenhouse emissions. Companies are increasingly aware of the negative effects of their operations on the environment and the need to reduce them [25]. Therefore, attention has also shifted towards the high energy consumption of data centres [9, 18, 21], pressuring them to reduce their carbon footprint.

Energy efficiency refers to solutions in which technology is used for the same functions with less energy [21]. Energy efficiency metrics can be used to evaluate data centre performance and find opportunities to reduce electricity use [9]. The high energy consumption and costs of data centres are mainly due to the electricity demand of the servers and the cooling system as well as lighting, electricity distribution and other support functions [19, 21]. Many data centre companies also use an excessive amount of energy to ensure high reliability for their customers, resulting in a low utilisation rate of IT systems [23].

2.4 Waste Heat

Data centres produce a significant amount of waste heat, typically quite low-grade, and it can be utilised in various applications [26]. Low heat can be raised in an economically profitable way; for example, with the help of heat pumps, it can be reused in a process that requires higher heat. Utilising waste heat can improve the energy efficiency of data centres by eliminating the need for chillers [18]. The temperature of waste heat depends on the cooling techniques used by the data centre [17].

With the help of heat pumps, waste heat can be used as a source of district heat [18]. For example, in Mäntsälä, an ecosystem has been built around the local data centre, where the waste heat is utilised in the municipality's district heating network [27, 28]. Stockholm City's district heating network also receives green energy by using waste heat from data centres. The ecosystem's vision is that no heat would be wasted in the data centre industry. Waste heat is therefore seen as a valuable resource, and its usage is part of an environmental act [29].

Waste heat can also be used for drying or heating in industrial processes [17]. Other examples of uses are greenhouses and fish farming. The Agriport A7 business ecosystem area is located near Amsterdam, where waste heat from the data centre is used in the operation of greenhouses [30]. According to Yuan et al. [31], the most efficient way to utilise waste heat is as a heating supply for nearby buildings.

3 RESEARCH METHODOLOGY

3.1 Research Process and Method

This study was carried out using a qualitative research method to increase understanding of the data centre industry and to form an understanding of issues that must be considered when creating a business ecosystem based on the utilisation of data centre waste heat. The empirical research included semi-structured interviews with experts and stakeholders in the case business ecosystem. Experts were interviewed to gain a deeper understanding of the industry.

The first interviewee was a data centre industry expert, and the second interviewee was a regional policy expert. The research was continued by interviewing the representatives of the companies in the case business ecosystem. The case business ecosystem companies are presented in Tab. 1. The data collected from the interviews were analysed using thematic analysis. Additional information about the data centre business ecosystems was gathered by sending a questionnaire to the following groups of stakeholders: 24 data centre companies, 72 potential user companies and 83 cities and municipalities. In total, 25 replies were received, making the combined response rate 14%. The questionnaires were intentionally kept concise and short to enable responses to be quick and easy. Most of the questions were multiple choice; for some, it was possible to select several different answer options, and some questions could be answered openly by writing a response. The answer options for the multiple-choice questions were 'yes', 'no' and 'I don't know'. The questionnaires were created using the Webropol tool, and a link to the surveys was sent to different target groups via email. The low number of responses can partly be explained by the fact that the recipients did not belong to the appropriate target group in terms of survey content. However, everyone who started the survey completed it, which suggests that the questions were not too difficult.

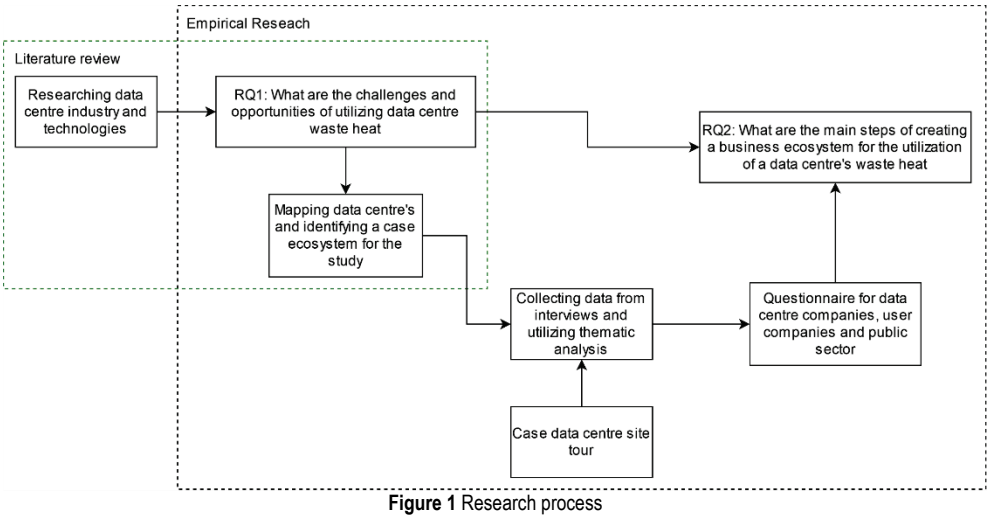


Figure 1 Research process

Table 1 Case ecosystem main players

Company	Description
A	Data centre company
B	District heating company
C	Material drying company

According to Flick [32], qualitative research focuses on collecting, interpreting and understanding qualitative data. Qualitative research can include many different interview formats, the most common of which is the semi-structured interview. Semi-structured interviews are flexible and allows the researcher to explore topics spontaneously during the interview [33]. Expert interviews can be used for various purposes, such as research data collection or familiarisation with a new field, to obtain a thematic structure and hypothesis for the researched topic [32]. According to Braun and Clarke [34], thematic analysis is a flexible analysis that is used to identify, analyse and finally report themes found in the data. Braun and Clarke [34] defined six steps for thematic analysis: familiarise yourself with the data you have collected, generate initial codes, find themes, review themes, define and name themes and produce a report.

3.2 Thematical Analysis

After coding the data from the case business ecosystem stakeholder interviews, the codes were divided into suitable themes and the themes were named descriptively. The themes and codes are presented in Tab. 2.

Table 2 Themes of thematical analysis

Theme	Codes
The business for waste heat utilisation	Ecosystem's stakeholders
	Roles in the ecosystem
	Legislation
Circular economy in the business ecosystem	Circular economy practices
	Environmental impacts
Lucrativeness of the business ecosystem	Profitability
	Quality of waste heat
	Challenges
	Benefits

3.2.1 Main Players of the Business Ecosystem for Waste Heat Utilisation

When creating a business model, all stakeholders of the ecosystem must be determined. In addition to the data centre and the user of waste heat (local heat or district heat), the

business ecosystem's stakeholders can include a city or municipality. The role of the city or municipality can be, for example, landowner or financier. From Company C's point of view, a significant factor affecting lucrativeness is also finding a customer in the data centre area who has material to be dried.

Company A's role in the business ecosystem would be mainly to be a passive donor of waste heat. According to them, they do not currently utilise waste or any heat recovering solutions, and see that the current electricity tax reduction for data centres has been designed with large operators in mind and therefore does not benefit smaller data centres. Company B, on the other hand, would operate and decide by itself when it is worth using waste heat in the district heating network and, if necessary, use some other heat source. In the case of Company C, the use of waste heat would depend on the model implemented in the project. One possibility would be to build a testing or pilot site that Company C itself or their customer would use independently, in which case the utilisation of waste heat would be irregular but still predictable.

Company B stated the following: 'Both the economic and environmental benefits of waste heat utilization are significant. Support through legislation can also be used to motivate district heating companies to make use of waste heat'. The attractiveness is increased by electricity tax reduction and constantly tightening regulations on fossil fuel use. From Company A's point of view, electricity tax reduction benefits larger data centres more than their operations.

3.2.2 Lucrativeness of the Case Business Ecosystem

Every stakeholder in the case project emphasised the impact of the profitability of the ecosystem on its lucrativeness. It was agreed that environmental benefits bring added value after financial benefits. In each company, however, the impact of environmental benefits on improving the company's image was identified.

Company A produces waste heat continuously around the clock and throughout the year. The amount of waste heat needed by both Companies B and C can vary at different times. The price of waste heat must be defined as appropriate and competitive for each party, and it must be predictable.

From Company A's point of view, building a business ecosystem for the utilisation of waste heat does not present significant challenges. From Company B's point of view, risks can arise from the operation of the heat pumps and whether they will last for at least the payback period of the investment. From Company C's point of view, the challenge of creating an ecosystem is the quality and quantity of waste heat, as the drying process requires a lot of thermal energy.

Company A would benefit financially from the sale of waste heat, as it would reduce cooling costs and be a new source of income. Companies B and C could also benefit financially from the ecosystem by gaining access to an affordable heat source. In addition, utilising emission-free waste heat could reduce the use of fossil fuels and thus increase the environmental friendliness of companies. Especially for Company B, the third significant benefit is improving the image of district heating through environmental friendliness.

4 RESULTS

4.1 Main Players of the Business Ecosystem

The concept was created based on the already existing data centre, which aims to collect and sell the waste heat in a business ecosystem and the corresponding literature. When building a business ecosystem, the stakeholders must be determined. The waste heat can be used in local or distant heating. Factors that fundamentally affect the choice of waste heat user are the location of the data centre and the amount and temperature of waste heat. In addition to the data centre and waste heat user, the city or municipality usually plays an important role as an enabler of the ecosystem. The roles of the stakeholders must also be determined because clear roles strengthen trust among stakeholders [35]. The identified prerequisites, challenges and opportunities of building a data centre waste heat business ecosystem are presented in Tab. 3.

Table 3 The prerequisites, challenges and opportunities of building a data centre waste heat business ecosystem

Prerequisites	Challenges	Opportunities
- Economic feasibility - Demand for waste heat - Stable and constant availability of waste heat - Desirable location of the data centre	- Payback of the investment - Predictability of electricity price development - Integration of waste heat recovery into the overall cooling system - Price and temperature of the waste heat - Balancing stakeholder requirements and responsibilities - Slow pace of zoning processes	- Reduction of cooling needs and costs - Reduction of burning more polluting sources for district heating, e.g. fossil fuels - Improvement of energy efficiency - Positive reputation - New source of income

Next, the rules, practices and business models for cooperation are determined. In addition, the contractual basis is agreed upon, which includes elements such as the price of the waste heat and it can be based, for example, on an open district principle. The challenges and benefits of the cooperation should also be recognised. The operations of the stakeholders should not be threatened because of the cooperation, and for example, the heat pumps must work at least for the payback time of the investment. The benefits of being part of the ecosystem can include cost savings, new incomes, energy efficiency improvements, reduction of emissions and improvement of the image of the companies.

The infrastructure for the ecosystem's operations must be planned. The infrastructure includes all the technology and equipment needed to recover and use the waste heat. These can include heat pumps and technology that is integrated into the data centre's cooling system. In addition, all required building and environmental permits must be applied for. Once the planning is done, the operation of the ecosystem can start. At this point, it must be confirmed that the recovery and usage of the waste heat are working efficiently. To ensure the sustainability of the ecosystem,

operations must also be continuously improved. The concept of the business ecosystem and the relationships between its parts are presented in Fig. 2.

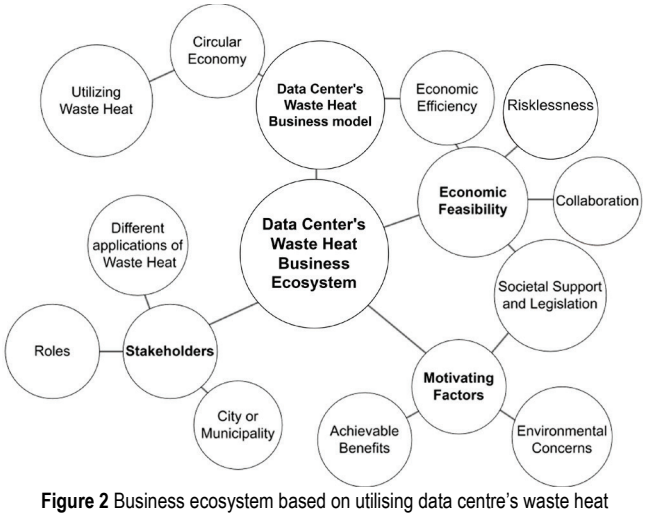


Figure 2 Business ecosystem based on utilising data centre's waste heat

The elements of the concept presented above offer the basis for building a business ecosystem where the data centre's waste heat is utilised. However, every project is unique, and the aspects of creating an ecosystem depend on the circumstances and stakeholders. When developing a business ecosystem, it is crucial to identify the stakeholders involved. Initially, the ecosystem area must have a user for the waste heat, either for the district or local heating. The city or municipality also plays a significant role as an enabler of the ecosystem. Additionally, the role of each stakeholder within the ecosystem must be defined clearly to enhance trust among stakeholders [35]. Generally, the data centre acts as a passive heat supplier while the waste heat user is the heat purchaser. The needs and relationships of the stakeholders are vital for value creation within the ecosystem. Belonging to the ecosystem emphasises cooperation, which requires transparency regarding the needs and expectations of each party. The compatibility of stakeholders' operations affects the ecosystem's functionality, and for the ecosystem to be sustainable, stakeholders must be committed to cooperation. The business ecosystem relies on sharing resources among stakeholders, where one company's waste is utilised in another company's value-creation process [13].

Next, a business model is established to define how costs, revenues and profits are generated from operations [36], and business models based on the circular economy have been developed to help companies transition to a circular economy. The development of the business model is influenced by the size of the entities involved, including the data centre and the waste heat user's heating needs. The utilisation of waste heat is based on a circular economy business model, which considers both economic value generation and environmental perspectives [37]. The circular economy is realised in the ecosystem by utilising the by-product of the data centre, namely waste heat. The profitability of an ecosystem primarily depends on its economic viability. Each stakeholder must benefit financially, which is facilitated by a well-structured contract and a competitive price for the waste heat. The contract can be based on principles such as bidirectional district heating or models defined by the energy authority. The economic benefit of using waste heat must cover all investment costs. A competitive price enables cost reductions for waste heat users and provides a new revenue stream for the data centre.

The safety and reliability of waste heat recovery and usage are essential for the ecosystem's profitability. The operational activities of the stakeholders must not be compromised; for example, heat pumps must function for the promised duration. Optimal utilisation of waste heat is continuous, but predictability in demand and pricing increases operational security. The suitability of waste heat usage must also consider specific applications. If necessary, the temperature of the waste heat can be increased using heat pumps to make it suitable for district heating. Local heating is the optimal option, as it avoids heat loss. Environmental concerns and the achievement of climate goals motivate the construction of an ecosystem. Circular economy principles can address sustainable development and climate goals [6]. Utilising waste heat can reduce emissions from combustion-based heat production. The growing energy demand and the desire to improve energy self-sufficiency also increase the viability of waste heat utilisation.

The international, national and private sectors all have significant roles in enabling sustainable development. Societal support can enhance the ecosystem's viability through legislation, such as potential reductions in electricity taxes, to encourage companies to utilise waste heat. An increasing number of European countries have either adopted, are currently adopting or have scheduled plans to phase out fossil fuel-based heating [38], making waste heat an ideal alternative heat source. Ecosystem operations can achieve environmental benefits through improved energy efficiency and reduced emissions. These environmental benefits can also enhance corporate image. However, one of the primary goals of participating in the ecosystem is economic gains, with environmental benefits as added value. According to the Ellen MacArthur Foundation [6], a circular economy can also lead to social benefits. Using waste heat can provide social benefits by supporting local communities, for example, by offering cheaper district heating to local residents.

4.2 Building a Data Centre Waste Heat Business Ecosystem

The starting point for the formation of a business ecosystem for waste heat involves the presence of an existing data centre generating waste heat that is not yet utilised. The formation of the ecosystem is significantly influenced by the location of the data centre [17], the quantity of waste heat produced and its temperature. The choice of waste heat user is influenced by the data centre's location, the amount of waste heat generated and its temperature. At this stage, it is necessary to determine the suitability of waste heat from the user's perspective. The process flow of building the data centre's waste heat business ecosystem is presented in Fig. 3 and is discussed in this section.

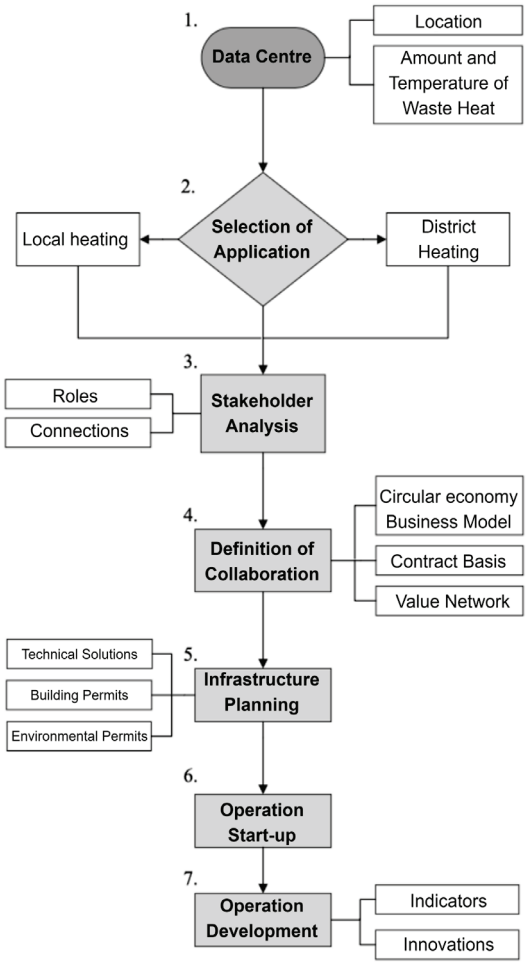


Figure 3 Building the main elements of the business ecosystem using the data centre's waste heat based on analysis

If there is vacant space around a data centre, local heating solutions, such as greenhouses, laundries or drying facilities for process industries, can be constructed. If the data centre is located in a residential area, the waste heat can be used to heat nearby buildings or streets. If the data centre is within a district heating network, the waste heat can be fed into it.

Next, the stakeholders of the ecosystem, their roles and their relationships are identified. Stakeholders include all parties who benefit from or influence the ecosystem's formation and development. These stakeholders include the data centre, the waste heat user, landowners, local residents and customers. Once the stakeholders are identified, their roles and responsibilities within the ecosystem are defined. In addition to clear roles that build trust among the ecosystem's stakeholders [35], establishing common rules are crucial for effective cooperation. According to Valkokari et al. [35], setting common rules can also enhance trust among the ecosystem's stakeholders. The needs and expectations of each actor must be understood, and interactions should be transparent. Additionally, the value network and business model of the ecosystem are defined, outlining how costs and revenues are structured within the ecosystem. Based on the business model, a contract framework is created that may follow principles such as bidirectional district heating or other existing contract templates. The framework specifies the price of the waste heat and other terms of utilisation. The price is determined based on investment and other costs, as well as the characteristics of the waste heat.

Once the business model and stakeholder needs are defined, the infrastructure of the ecosystem must be planned. This includes all necessary technologies and equipment for capturing and utilising waste heat, such as potential heat pumps and technology integrated into the data centre's cooling system. Additionally, the required environmental and building permits must be obtained. After the ecosystem's design is complete and all necessary infrastructure is built, operations can commence. At this stage, the operational activities of the ecosystem are initiated, ensuring efficient heat recovery and utilisation. To ensure the ecosystem's sustainability and continued operation, continuous development is necessary. According to Moore [39], the existence of a business ecosystem depends on its renewal in response to external changes, such as innovations or technologies. The ecosystem can be developed through activities such as operational optimisation and the creation of innovations. Development areas can be identified using various metrics, such as power usage effectiveness (PUE) and energy reuse factor (ERF), to measure the data centre's energy efficiency.

These phases provide a foundation for constructing a circular economy-based business ecosystem that utilises waste heat from data centres. However, it is important to note that every project is unique and that the phases of ecosystem development may vary depending on a project's conditions and different stakeholders. Additionally, the process of creating a new ecosystem inherently involves zoning and building permit processes.

4.3 Practical Implications

The concept of a business ecosystem based on the utilisation of waste heat was created at the general level from the point of view of an already existing data centre. The main pillars of the concept are the different options for the use of waste heat as well as the profitability, challenges and benefits of the ecosystem. Regarding the research questions, challenges for waste heat utilisation were detected to be payback of the investment, predictability of electricity price development, integration of the waste heat recovery into the overall cooling system, price and temperature of the waste heat, balancing stakeholder requirements and responsibilities and the slow pace of zoning processes. The opportunities detected were the reduction of cooling needs and costs, reduction of burning more polluting sources for district heating (e.g. fossil fuels), improvement of energy efficiency, positive reputation and a new source of income. The main steps of building a data centre ecosystem are the following: specifying the requirements of the data centre, selecting the application, conducting stakeholder and collaboration model analyses, planning the infrastructure, followed by the operation start-up and development.

Factors affecting profitability are the cost and earnings logic of the ecosystem's business model, value creation and the characteristics of waste heat. A separate business model must be built for each project, depending on the size, needs and contract basis of the operators, among other things. In

particular, the size of the data centre and the amount of waste heat impact the business ecosystem that can be built around it.

The selection of the waste heat user is influenced, among other things, by the location of the data centre and the amount and quality of the waste heat. The users of waste heat can be roughly divided into local and district heating. The primary use is heating office premises in nearby areas. If enough waste heat is generated, it makes sense to use it, for example, in district heating or in heat-demanding processes such as local heating in the data centre area. Waste heat in district heating has been widely implemented, although this approach has been questioned in recent years due to energy losses in district heating pipelines. Thermal value is also influenced by other factors, such as soil conditions, infrastructure, insulation and the depth and construction of the district heating network.

Data centres would benefit economically from the ecosystem because it would serve as a new source of income, and customers would acquire a budget source of environmentally friendly heat. Currently, combustion based on biomass and other renewable energy sources is still profitable, but this will change in the near future. Therefore, the long-term sustainability of combustion-based energy production is poor. With the green transition, the first step is to abandon the burning of fossil fuels in favour of more environmentally friendly fuels and ultimately to reduce combustion altogether by utilising, among other things, the waste heat from data centres.

5 CONCLUSIONS

The purpose of this study was to create a concept for a business ecosystem that utilises the waste heat produced by a data centre in other companies' operations that require heat. One of the driving forces behind the promotion of waste heat utilisation in data centres is the requirements set by the EU directive on waste heat recovery [7]. The starting point of the concept is an already existing data centre creating waste heat around which a business ecosystem is planned.

When considering an ecosystem for data centre waste heat utilisation, key aspects include exploring different user options and assessing profitability, challenges and benefits. Factors affecting profitability include cost, value creation and quality of the waste heat. The business model should be tailored to each project based on the data centre's size and needs. Selecting a waste heat user depends on location. A specific business model defines costs, revenue and value creation, and clear contractual agreements are essential. Stakeholders in the ecosystem are identified, encompassing all factors influencing waste heat utilisation.

Key challenges in utilising waste heat from data centres include investment payback, electricity price uncertainty, technical integration, heat quality, stakeholder coordination and zoning delays. However, opportunities lie in reduced cooling costs, lower fossil fuel use in district heating, improved energy efficiency, enhanced reputation and new revenue streams. The formation of a business ecosystem for data centre waste heat recovery begins with assessing the location, quantity and temperature of available waste heat. Based on these, local or district heating are suitable applications. Key phases include identifying stakeholders and their roles, defining collaboration models and business contracts, planning infrastructure and obtaining permits and starting operations. Continuous development through monitoring, optimisation and innovation ensures the ecosystem remains sustainable and adaptive.

The research could be expanded by validating the results quantitatively and by repeating the study with a different ecosystem in hand. This research was based on a case study and results may vary based on the regional attributes and are also affected by the number of interviews and respondents. A similar concept could also be built from the perspective of a new data centre. In such a case, the same issues should be considered, but regional planning would represent a bigger factor. In addition, further studies could focus on how waste heat and other side streams can be utilised in larger-scale circular business ecosystems, such as in the steel industry. For example, an interesting and relevant question in this field is the decarbonisation of the steel industry, and one way of contributing to that could be by producing waste heat to produce biochar pellets to substitute fossil coal in electric arc furnaces.

6 REFERENCES

[1] Urbinati, A., Chiaroni, D., & Chiesa, V. (2017). Towards a new taxonomy of circular economy business models. *Journal of Cleaner Production, 168*, 487–498. https://doi.org/10.1016/j.jclepro.2017.09.047

[2] Murray, A., Skene, K., & Haynes, K. (2017). The circular economy: An interdisciplinary exploration of the concept and application in a global context. *Journal of Business Ethics, 140*, 369–380. https://doi.org/10.1007/s10551-015-2693-2

[3] Schulte, U. G. (2013). New business models for a radical change in resource efficiency. *Environmental Innovation and Societal Transitions, 9*, 43–47. https://doi.org/10.1016/j.eist.2013.09.006

[4] Okorie, O., Salonitis, K., Charnley, F., Moreno, M., Turner, C., & Tiwari, A. (2018). Digitisation and the circular economy: A review of current research and future trends. *Energies, 11*(11). https://doi.org/10.3390/en11113009

[5] Geissdoerfer, M., Savaget, P., Bocken, N. M., & Hultink, E. J. (2017). The circular economy–A new sustainability paradigm? *Journal of Cleaner Production, 143*, 757–768. https://doi.org/10.1016/j.jclepro.2016.12.048

[6] Ellen MacArthur Foundation. (2019). Completing the picture: How the circular economy tackles climate change.

[7] EU. (2023) Directive (EU) 2023/1791 of the European Parliament and of the Council of 13 September 2023 on energy efficiency and amending regulation (EU) 2023/955 https://eur-lex.europa.eu/eli/dir/2023/1791/oj/eng

[8] Leppänen, T., Romka, R., & Tervonen, P. (2020). Utilization of data centre waste heat in Northern Ostrobothnia. Tehnički Glasnik, *14*(3), 312–317. https://doi.org/10.31803/tg-20200706172649

[9] Ni, J., & Bai, X. (2017). A review of air conditioning energy performance in data centers. *Renewable and Sustainable Energy Reviews, 67*, 625–640. https://doi.org/10.1016/j.rser.2016.09.050

[10] Khalaj, A. H., Scherer, T., & Halgamuge, S. K. (2016). Energy, environmental and economical saving potential of data centers with various economizers across Australia. *Applied Energy, 183*, 1528–1549. https://doi.org/10.1016/j.apenergy.2016.09.053

[11] IEA. (2022). Data centres and data transmission networks, IEA, Paris https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks#tracking
[12] EU. (2008). The European Parliament and of the Council on waste and repealing certain directives 2008/98/EY, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32008L0098
[13] Roos, G. (2014). Business model innovation to create and capture resource value in future circular material chains. *Resources, 3*(1), 248–274. https://doi.org/10.3390/resources3010248
[14] Leppänen, T., Mustonen, E., Saarela, H., Kuokkanen, M., & Tervonen, P. (2020). Productization of industrial side streams into by-products—case: Fiber sludge from pulp and paper industry. *Journal of Open Innovation: Technology, Market, and Complexity, 6*(4), 185. https://doi.org/10.3390/joitmc6040185
[15] Husgafvel, R., Karjalainen, E., Linkosalmi, L., & Dahl, O. (2016). Recycling industrial residue streams into a potential new symbiosis product–The case of soil amelioration granules. *Journal of Cleaner Production, 135*, 90–96. https://doi.org/10.1016/j.jclepro.2016.06.092
[16] Lee, H., & Schmidt, G. (2016). Using value chains to enhance innovation. *Production and Operations Management, 26*. https://doi.org/10.1111/poms.12665
[17] Koronen, C., Åhman, M., & Nilsson, L. J. (2020). Data centres in future European energy systems—Energy efficiency, integration and policy. *Energy Efficiency, 13*(1), 129–144. https://doi.org/10.1007/s12053-019-09833-8
[18] Oró, E., Depoorter, V., Garcia, A., & Salom, J. (2015). Energy efficiency and renewable energy integration in data centres. Strategies and modelling review. *Renewable and Sustainable Energy Reviews, 42*, 429–445. https://doi.org/10.1016/j.rser.2014.10.035
[19] Rong, H., Zhang, H., Xiao, S., Li, C., & Hu, C. (2016). Optimizing energy consumption for data centers. *Renewable and Sustainable Energy Reviews, 58*, 674–691. https://doi.org/10.1016/j.rser.2015.12.283
[20] Avgerinou, M., Bertoldi, P., & Castellazzi, L. (2017). Trends in data centre energy consumption under the European code of conduct for data centre energy efficiency. *Energies, 10*(10), 1470. https://doi.org/10.3390/en10101470
[21] Uddin, M., & Rahman, A. A. (2012). Energy efficiency and low carbon enabler green IT framework for data centers considering green metrics. *Renewable and Sustainable Energy Reviews, 16*(6), 4078–4094. https://doi.org/10.1016/j.rser.2012.03.014
[22] Daim, T., Justice, J., Krampits, M., Letts, M., Subramanian, G., & Thirumalai, M. (2009). Data center metrics: An energy efficiency model for information technology managers. *Management of Environmental Quality: An International Journal, 20*(6), 712–731. https://doi.org/10.1108/14777830910990870
[23] Bertoldi, P. (2014). A market transformation programme for improving energy efficiency in data centres. In ACEEE Summer Study on Energy Efficiency in Buildings (pp. 9–14).
[24] Sovacool, B. K., Monyei, C. G., & Upham, P. (2022). Making the internet globally sustainable: Technical and policy options for improved energy management, governance and community acceptance of Nordic datacenters. *Renewable and Sustainable Energy Reviews, 154*, 111793. https://doi.org/10.1016/j.rser.2021.111793
[25] European Commission. (2016). EU code of conduct on data centre energy efficiency – Participant guidelines. Ispra, Version 3.0.0.
[26] Acton, M., Bertoldi, P., & Booth, J. (2021). Best practice guidelines for the EU Code of conduct on data centre energy efficiency, European Commission. Ispra, 2021, JRC123653.
[27] Nivos. (2016). Paikallisen datakeskuksen kanssa yhteistyössä luotu ekosysteemi lämmittää Mäntsälässä. https://nivos.fi/paikallisen-datakeskuksen-kanssa-yhteistyossa-luotu-ekosysteemi-lammittaa-mantsalassa/ (Accessed 30.3.2023).
[28] Mäntsälän yrityskehitys. Mäntsälän ekosysteemi – Oletko valmis kiertotalouteen? https://www.yrityskehitys.net/mantsalan-ekosysteemi (Accessed 30.3.2023).
[29] Stockholm Data Parks. https://stockholmdataparks.com/green-computingredefined/ (Accessed 30.3.2023).
[30] Agriport A7. Agriport datacenter area. https://www.agriporta7.nl/agriportdatacenter-area/?lang=en (Access 4.4.2023].
[31] Yuan, X., Liang, Y., Hu, X., Xu, Y., Chen, Y., & Kosonen, R. (2023) Waste heat recoveries in data centers: A review, *Renewable and Sustainable Energy Reviews*, *188*. https://doi.org/10.1016/j.rser.2023.113777
[32] Flick, U. (2009). An introduction to qualitative research (4th ed.). SAGE.
[33] Ryan, F., Coughlan, M., & Cronin, P. (2009). Interviewing in qualitative research: The one-to-one interview. *International Journal of Therapy and Rehabilitation, 16*(6), 309–314. https://doi.org/10.12968/ijtr.2009.16.6.42433
[34] Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
[35] Valkokari, K., Hyytinen, K., Kutinlahti, P., & Hjelt, M. (2020). Yhdessä kestävää kasvua -ekosysteemiopas. VTT Technical Research Centre of Finland.
[36] Teece, D. J. (2010). Business models, business strategy and innovation. *Long Range Planning, 43*(2–3), 172–194. https://doi.org/10.1016/j.lrp.2009.07.003
[37] Adams, R., Jeanrenaud, S., Bessant, J., Denyer, D., & Overy, P. (2016). Sustainability-oriented innovation: A systematic review. *International Journal of Management Reviews, 18*(2), 180–205. https://doi.org/10.1111/ijmr.12068
[38] Braungardt, S., Tezak, B., Rosenow, J., & Bürger, V. (2023). Banning boilers: An analysis of existing regulations to phase out fossil fuel heating in the EU. *Renewable and Sustainable Energy Reviews*, *183*. https://doi.org/10.1016/j.rser.2023.113442
[39] Moore, J. F. (1993). Predators and prey: A new ecology of competition. *Harvard Business Review, 71*(3), 75–86.

Authors' contact details:

Juuso Kivijakola, MSc
(Corresponding author)
University of Oulu,
Pentti Kaiteran katu 1, FI-90014 Oulu, Finland
juuso.kivijakola@oulu.fi

Silja Heikkinen, MSc.
University of Oulu,
Pentti Kaiteran katu 1, FI-90014 Oulu, Finland
silja.heikkinen@oulu.fi

Pekka Tervonen, DSc
University of Oulu,
Pentti Kaiteran katu 1, FI-90014 Oulu, Finland
pekka.tervonen@oulu.fi

Harri Haapasalo, Prof., DSc
University of Oulu,
Pentti Kaiteran katu 1, FI-90014 Oulu, Finland
harri.haapasalo@oulu.fi

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240719121508

Preliminary communication
Received: 2024-07-19, Accepted: 2025-01-20

Tensile Properties Regression Analysis of FDM 3D Printed Composites with ONYX Matrix and Different Fibres

Pejo Konjatić, Josip Kačmarčik*, Marko Katinić, Lamija Mešeljević

Abstract: The influence of infill density, number of layers, and fibre type on the tensile mechanical properties of composite parts produced by FDM 3D printing with continuous fibre reinforcement (CFR) was investigated. Specimens made of ONYX composite material reinforced with carbon, glass, and aramid fibres were tested using a static tensile test according to ISO 527. A linear regression model was developed to correlate mechanical properties with 3D printing parameters. However, the influence of infill density could not be reliably determined due to the automatic generation of solid infill around fibre layers and at boundary layers in the utilized 3D printing software. These reinforcements provided varying degrees of enhancement, with carbon and glass fibres showing the highest increase in strength, glass fibres offering the best enhancement in fracture strain, and carbon fibres in stiffness. The obtained models for tensile strength, strain at maximum stress, and modulus of elasticity can be useful in the design of 3D printed parts, offering a simple solution for the prediction of mechanical properties.

Keywords: 3D printing; composite materials; aramid fibres; carbon fibres; fibreglass; tensile properties

1 INTRODUCTION

Additive manufacturing, commonly known as 3D printing (3DP), revolutionized the process of design and creation in various industries [1-3]. The most accessible, economical, and widely applicable 3D printing technology in recent years is FDM (Fused Deposition Modelling), where the process is based on the extrusion and melting of thermoplastic material (filament) through a heated nozzle and layering it onto a platform along a predefined path, where the filament cools and solidifies to form a 3D object. The main ad-vantage of selecting the FDM process is the production of components without the need for pre-prepared tools, such as injection moulds, making it very useful for rapid prototyping. Since FDM 3D printing is a relatively new technology, the mechanical properties of the materials produced and their testing still present many unknowns and have been the subject of numerous recent studies [4-8].

FDM printing typically uses thermoplastic polymers like polylactic acid (PLA), acrylonitrile butadiene styrene (ABS), and polyethylene terephthalate glycol (PETG), and nylon, each offering different levels of strength, quality, and cost. Recent advancements include materials reinforced with chopped microfibres, such as carbon or glass, which significantly enhance mechanical properties and thermal stability. This range of materials allows for greater flexibility in selecting the appropriate material for specific applications. However, despite recent advancements, conventional FDM processes still face challenges such as lower strength and stiffness, poor dimensional accuracy and surface finish, and limited thermal stability compared to traditional manufacturing methods and compo-sites. Development of technology for 3DP of continuous fibre reinforced addressed these shortcomings and make a milestone in applicability and reliability of 3D printed parts [9, 10]. Unlike the chopped microfibres, the continuous fibres are printed directly into the parts, serving as the structural foundation and enhancing overall performance offering better strength, uniformity, and enhanced structural integrity. The technology available in the market with this feature is also known as continuous filament fabrication (CFF). However, 3DP is a challenging process where different process parameters influence the result. This has prompted numerous studies in recent years on the impact of various composite 3DP process parameters on different characteristics of the final part. The improvement of mechanical properties is one of the most important benefits of 3DP composites, so a significant number of research studies address it. Tensile testing is a fundamental and widely used test for examining the mechanical properties of various materials, providing essential information about the material's mechanical behaviour necessary for designers and engineers, so it is also widely implemented in 3DP research.

Analysing the influence of matrix and fibre material for FDM printing filament, process parameters and developing analytical/numerical models for predicting the mechanical properties of 3DP is one of the important research directions. The ability of known micromechanical models to predict modulus of elasticity and tensile strength for 3D printed composites with different glass fibre volume fraction is investigated in [11]. In the [12] the effect of various FDM printing parameters on the impact resistance of 3D-printed composites with glass fibre is investigated. Similarly, in [13], the dependence of the modulus and the strength on concentric carbon fibre reinforcement and density of 3D printed materials is examined, and models for optimization of 3DP configuration are proposed. Regression analysis of tensile strength, modulus and energy absorption as function of 3DP parameters is conducted in [14]. The effect of the material storage and aging is also investigated [15]. Interested readers can find numerous other studies of similar content that further underscore the significance and appeal of this topic.

The influence of fill density and number of layers with fibre reinforcement on tensile mechanical properties of 3D printed composites with ONYX material is experimentally and statistically investigated in this study. Additionally, three different types of fibre reinforcement, carbon, fibreglass and

aramid (Kevlar®) are considered. The ISO 527 standard [16-17], which defines the tensile testing of plastics, was applied for the investigation. Three characteristics were determined in the testing: tensile strength (maximum stress), elongation at tensile strength, and modulus of elasticity under tension (Young's modulus). Markforged Mark Two 3D printer [18] and materials for additive manufacturing from the Markforged [19] are used in experimental investigations. The aim of the research is to develop regression models that correlate various tensile mechanical properties with the parameters of the 3D printing process and usage of different fibres. Such models for different fibres may not be readily available in the open access literature. While modest in novelty, this work contributes to a deeper understanding of the relationships between process parameters and final mechanical properties in 3D printed composites. These models may be useful in optimizing the design and production of 3D prints, providing valuable guidance to the industry.

2 MATERIALS AND METHODS

2.1 Specimens Geometry

The ISO 527 standard regulates the test parameters for materials tested at room temperature. It defines the procedure for tensile testing of plastics, providing guidance on measuring mechanical properties such as tensile strength and elongation at break. The standard offers both normal (type 1) and small (type 5) specimens. For the convenience of faster 3D printing and compatibility with the test equipment regarding maximum force, small type 5A specimens were selected for investigation. The dimensions of the specimen are provided in Tab. 1 and Fig. 1.

Table 1 Dimensions of test specimen type 5A (ISO 527)

Feature	Dimensions / mm
l_2	≥ 75
b_2	12,5 ± 1
l_1	25 ± 1
b_1	4 ± 0,1
r_1	8 ± 0,5
r_2	12,5 ± 1
L	50 ± 2
L_0	20 ± 0,5
h	≥ 2

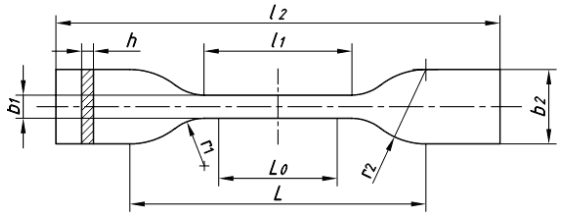


Figure 1 Tensile test specimen type 5A (ISO 527)

2.2 3D Printing

The Markforged Mark Two printer, as shown in Fig. 2, was used for 3D printing. It features a single print head equipped with two distinct nozzles. One nozzle is specifically designated for printing polymer material, while the other is dedicated to laying continuous fibre lines within the polymer. These fibres serve to reinforce the 3D printed parts, significantly enhancing their stiffness and strength. The FDM process proceeds continuously to produce a basic block of the 3D printed part, with the print head operating one nozzle at a time to avoid interaction between the two types of filament. This additive manufacturing process, known as Continuous Fibre Reinforcement (CFR), enhances the printed part with continuous fibres within the polymer matrix.

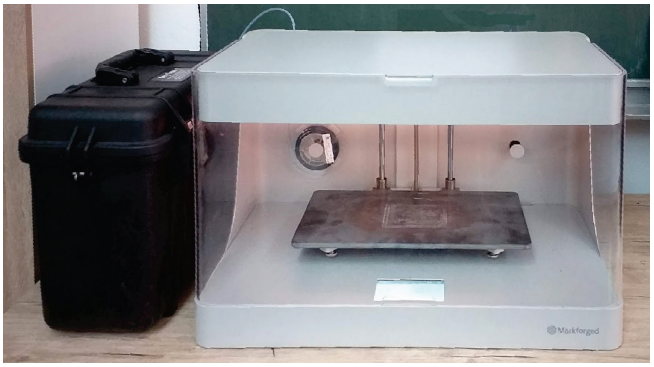


Figure 2 Markforged Mark Two 3D printer

The investigated specimens were 3D printed using ONYX as a composite base, rein-forced with continuous fibres from carbon, fibreglass, or aramid (Kevlar®), printed into the samples parallel to the ONYX printing process. ONYX is a relatively new material, a filled plastic composed of engineering nylon fused with chopped carbon micro fibres. It offers high strength, toughness, and chemical resistance, both as a standalone material and when reinforced with continuous fibres to produce aluminium-strength parts. Each type of utilized fibre possesses unique characteristics: carbon fibres excel in strength and rigidity, glass fibres provide a balance of strength and flexibility, and Kevlar fibres offer high strength combined with flexibility. The selection of fibres for composite 3D printing depends on specific application requirements and manufacturing preferences. Tab. 2 provides the mechanical properties of the materials utilized for 3D printing, as provided by the manufacturer (Markforged).

Table 2 Mechanical and physical properties of used materials [20]

Properties	Onyx	Carbon Fibre	Fibreglass	Kevlar
Density / g/cm^3	1,2	1,4	1,5	1,2
Tensile Stress at Yield / MPa	40	/	/	/
Tensile Strength / MPa	/	800	590	610
Tensile Modulus / GPa	2,4	60	21	27
Flexural Strength/ MPa	71	540	200	240
Tensile Strain at Break / %	25	1,5	3,8	2,7
Flexural Modulus / GPa	3	51	22	26

The Eiger software [21] is utilized for setting 3DP parameters and serves as the official 3DP software for all Markforged printers, providing users with control over the 3D printing process. Fig. 3 illustrates the setup of the 3D model of the specimen in Eiger software showing the position of the specimen on the printer bed for 3DP. The specimen 3D model is previously prepared in CAD software.

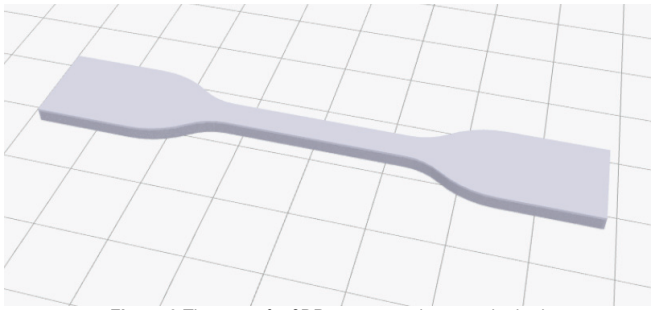


Figure 3 The setup for 3DP: a test specimen on the bed

In this study, the triangular fill pattern, which is widely used in 3D printing, is selected for the 3D printed composite matrix. For this pattern, adjustments of filling density ranging from 28% to 54% are provided in the Eiger software. Besides this minimum and maximum fill density values, the median value of 41% is investigated. A single line of material is designated as the wall thickness for parts. Default setup in the software, consisting of 4 solid plastic layers at the top and bottom, was retained for 3D printing the specimens. The concentric fibre fill type, selected from the options for reinforcing fibre layers in parts, was utilized. Additionally, the option of a single line of fibre placed adjacent to walls was selected. This resulted in two longitudinally oriented fibres, next to specimens' walls, in a reinforced layer in the narrow part of specimen relevant for tensile testing. An example of a reinforced layer of a test specimen, printed with 41% infill, is shown in Fig. 4. The number of reinforcement layers is varied, with choices ranging from 4, 6, to 8 layers. The layer height for 3DP process was 0,125 mm for specimens with carbon fibre, and 0,1 mm for fibreglass and Kevlar fibres reinforcement, corresponding to fibre thickness.

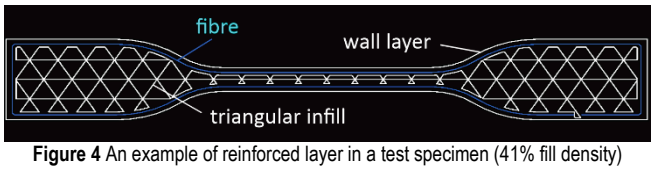


Figure 4 An example of reinforced layer in a test specimen (41% fill density)

2.3 Design of Experiment

Two numeric factors (fill density, number of reinforced layers) and one categoric factor (the different fibre type) were investigated in the research. A full factorial design with three levels of numeric factors is employed in the design of the experiment, with two additional central points, resulting in a total of 11 experimental points. This design is replicated for each of the three fibre types, yielding a grand total of 33 experimental points. The central points were used to check the linearity of the experimental system and to assess the impact of the factor means on the studied responses. Tab. 3 provides the plan of experiment, detailing the order and values of the experimental points. Additionally, Fig. 5 illustrates the implemented design of experiment with displayed coded and true values.

Table 3 Plan of experiment

Specimen	A: Fill Density in %	B: No. of reinforced layers	C: Fibre type
1	28	4	Carbon Fibre
2	54	4	Carbon Fibre
3	28	8	Carbon Fibre
4	54	8	Carbon Fibre
5	28	6	Carbon Fibre
6	54	6	Carbon Fibre
7	41	4	Carbon Fibre
8	41	8	Carbon Fibre
9	41	6	Carbon Fibre
10	41	6	Carbon Fibre
11	41	6	Carbon Fibre
12	28	4	Kevlar
13	54	4	Kevlar
14	28	8	Kevlar
15	54	8	Kevlar
16	28	6	Kevlar
17	54	6	Kevlar
18	41	4	Kevlar
19	41	8	Kevlar
20	41	6	Kevlar
21	41	6	Kevlar
22	41	6	Kevlar
23	28	4	Fibreglass
24	54	4	Fibreglass
25	28	8	Fibreglass
26	54	8	Fibreglass
27	28	6	Fibreglass
28	54	6	Fibreglass
29	41	4	Fibreglass
30	41	8	Fibreglass
31	41	6	Fibreglass
32	41	6	Fibreglass
33	41	6	Fibreglass

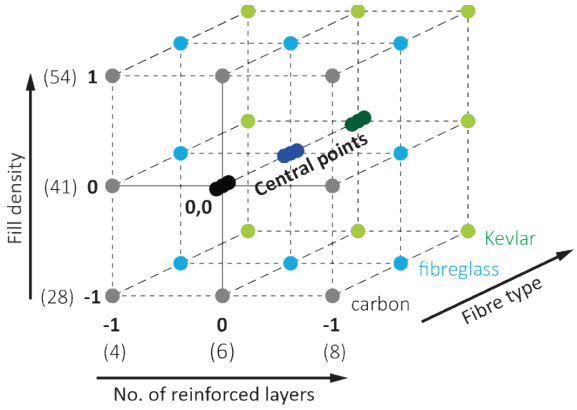


Figure 5 Three level full factorial design with two numeric and one categoric factor

In this study, explicit blocking was not incorporated into the experimental design. The experiment followed a full factorial design with replication and randomization, which allowed us to investigate the effects of the primary factors—fill density, number of reinforced layers, and fibre type—independently. Potential sources of variability, such as material batch differences or 3D printer calibration inconsistencies, were controlled by using the same material batch for all specimens and maintaining consistent calibration and environmental conditions throughout the study. These measures ensured homogeneity across the experimental units and reduced the need for blocking.

In order to increase the reliability of the results, randomization was integrated into the test procedure. The order of the samples was randomized to randomly distribute any uncontrolled variability across the test series.

3 RESULTS

3.1 Tensile Test Results

The test samples were tested at room temperature at a speed of 2 mm/min on a Shimadzu AGS-X 10 kN tensile testing machine (Fig. 6). Each test specimen was marked before testing to avoid potential errors due to accidental sample mix-up. An example of tested specimens for each fibre type is shown in Fig. 7.



Figure 6 Shimadzu AGS-X 10 kN tensile testing machine

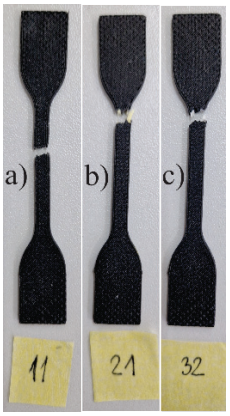


Figure 7 Examples of tested specimens with different fibre types: a) carbon, b) aramid, c) fibreglass

Using the data obtained from the static tensile test machine, the values for maximum stress, modulus of elasticity and deformation were calculated according to the ISO 527, and the obtained results are shown in the Tab. 4. Graphical representation of results is shown in charts on Fig. 8a, 8b and 8c, for maximum tensile stress (tensile strength), strain at maximum stress and tensile elastic modulus (Young's modulus), respectively.

Table 4 Tensile test results

Specimen	Maximum stress R_m / MPa	Strain at maximum stress ε_m	Young's modulus E / MPa
1	97,29	0,0610	2172,8
2	93,99	0,0719	2944,6
3	127,61	0,0632	3086,3
4	131,29	0,0527	2693,8
5	103,31	0,0664	2637,4
6	110,51	0,0578	2129,9
7	90,55	0,0682	2815,5
8	121,56	0,0673	2399,9
9	108,84	0,0578	2539,7
10	102,21	0,0593	2539,6
11	105,04	0,0934	1065,2
12	76,46	0,0851	1266,4
13	73,09	0,1014	1596,5
14	106,60	0,1018	1559,3
15	107,94	0,0907	1443,2
16	92,02	0,0975	1222,8
17	87,29	0,1031	1206,7
18	76,51	0,1120	1510,0
19	112,65	0,1001	1233,4
20	90,64	0,0981	1280,9
21	87,72	0,0974	1274,8
22	86,15	0,1360	1158,7
23	85,16	0,1273	1077,6
24	81,42	0,1360	1559,8
25	127,73	0,1732	1509,9
26	131,41	0,1554	1377,0
27	112,44	0,1498	1276,8
28	100,21	0,1360	1135,4
29	88,42	0,1188	1605,1
30	121,87	0,1423	1387,8
31	117,14	0,1349	1378,3
32	108,41	0,1343	1416,4
33	108,97	0,0610	2172,8

3.2 Regression Analysis

Regression analysis is done using Stat-Ease 360 [22], the commercial software with capabilities for data analysis, including tools for design and analysis of deterministic computer experiments. Stat-Ease automatically fits polynomial models up to sixth-order polynomials to model experimental results and afterwards based on statistical data such as *p*-values, lack of fit, and *R*-squared values compares the models, and recommends the statistically most appropriate model. The statistical data of regression analysis for suggested models of three responses in investigated experiment is given in Tab. 5 and Tab. 6.

The adjusted *R*-squared value considers the number of predictors in the model, while the predicted *R*-squared value estimates the model's performance on new data. Higher *R*-squared values indicate a better fit of the regression model to the data, implying that the model can effectively explain the

variability observed in the response variable. For example, the *R*-squared values for the response variable *E* (Young's modulus) indicate that approximately 97,16% of the variance is explained by the linear regression model, making it the most effective predictor among the three responses. Therefore, *E* is considered to provide the best results in this analysis.

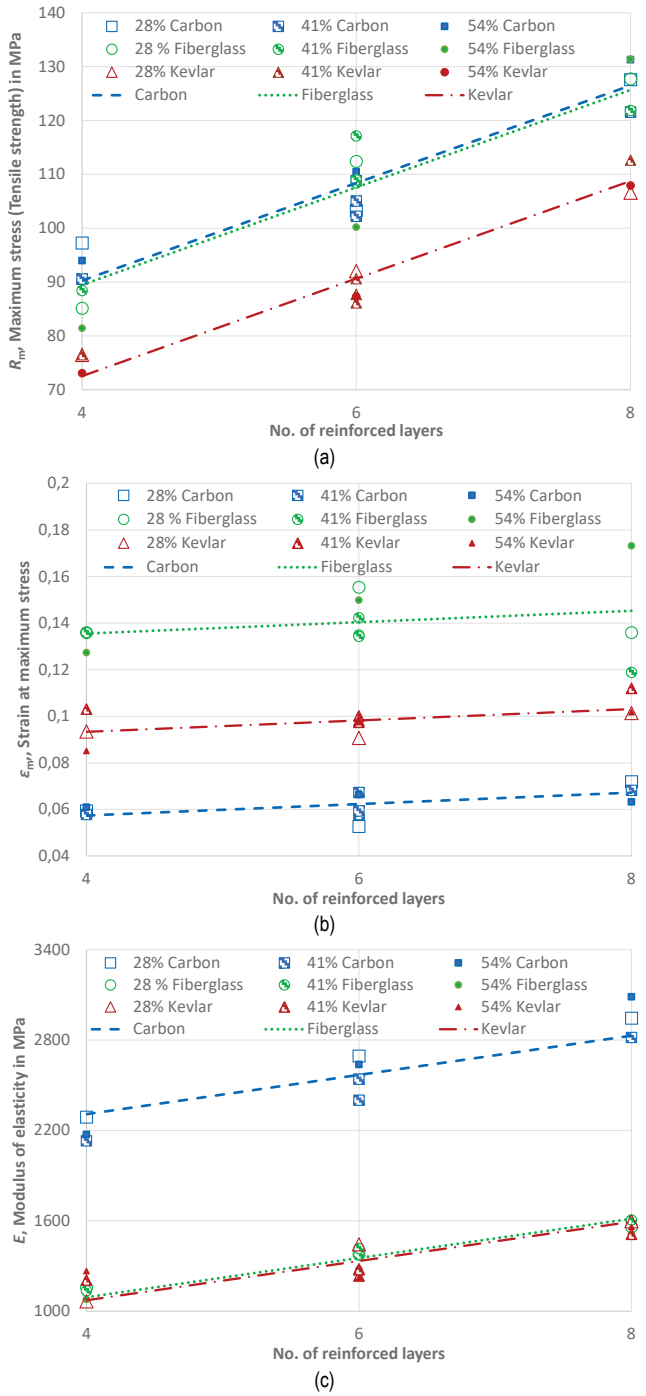


Figure 8 Tensile test results and linear regression models for (a) maximum stress (tensile strength), (b) strain at maximum stress, (c) Young's modulus

Table 5 Summary of statistical fit data

Response	Model	Model *p*-value	Lack of fit *p*-value	Lack of fit *F*-value	Adjusted R^2	Predicted R^2
R_m	Linear	$< 10^{-4}$	0,2593	1,72	0,923	0,9055
ε_m	Linear	$< 10^{-4}$	0,012	6,86	0,9218	0,901
E	Linear	$< 10^{-4}$	0,0227	5,34	0,9716	0,964

Table 6 ANOVA analysis for linear models' factors.

Response	Data	A: Fill Density	B: Reinforced layers	C: Fibre type
R_m	*F*-value	0,3493	281,73	52,7
	p-value	0,5593	$< 10^{-4}$	$< 10^{-4}$
ε_m	*F*-value	0,4995	4,74	188,11
	p-value	0,4856	0,038	$< 10^{-4}$
E	*F*-value	0,5016	109,63	494,61
	p-value	0,4846	$< 10^{-4}$	$< 10^{-4}$

The *p*-values less than 0,0500 indicate model terms are significant. In this case, the ANOVA analysis confirmed observation from the test results that A: Fill density factor is not significant. Drawing from that the modified linear models are generated in software, without the term for A: Fill density.

The inclusion of central points in the experimental design confirmed the linearity of the response surface, as no significant deviations were observed. This validated the use of linear regression models for the analysis, as demonstrated by the R^2 values presented in Tab. 5. The linear models for maximum stress, strain at maximum stress and modulus of elasticity obtained at the end of analysis, for three investigated fibre types, are shown in Fig. 8, alongside the corresponding experimental results for comparison. The coded equations for the models are:

$$R_m = 102{,}20 + 18{,}10 \cdot B + 6{,}19 \cdot C[1] + 5{,}37 \cdot C[2] \quad (1)$$

$$\varepsilon_m = 0{,}1003 + 0{,}0049 \cdot B - 0{,}038 \cdot C[1] + 0{,}0401 \cdot C[2] \quad (2)$$

$$E = 1751{,}19 + 260{,}35 \cdot B + 816{,}74 \cdot C[1] - 398{,}21 \cdot C[2] \quad (3)$$

The coded values of the factors in Eqs. (1) – (3) are given in Tab. 7. Coded values were used to enable the development of a single model encompassing all fibre types, as categorical variables can only be incorporated into the regression analysis using such coded values. For consistency, all other variables were also presented using coded values.

Table 7 Coded values for model factors

Numeric factors				Categoric factors		
Fill Density in %	A	Reinforced Layers	B	Fibre Type	C[1]	C[2]
28	–1	4	–1	Carbon	1	0
41	0	6	0	Fibreglass	0	1
54	1	8	1	Kevlar	–1	–1

ANOVA values for the modified models, excluding the fill density term, are given in Tab. 8. This exclusion resulted in models that better fit the experimental data, as evidenced by the increased F-values for the remaining factors, which indicates their stronger influence on the observed responses.

Table 8 ANOVA analysis for modified linear models' factors

Response	Data	B: Reinforced layers	C: Fibre type
R_m	*F*-value	288,2	53,91
	p-value	$< 10^{-4}$	$< 10^{-4}$
ε_m	*F*-value	4,83	191,41
	p-value	0,0362	$< 10^{-4}$
E	*F*-value	111,55	503,26
	p-value	$< 10^{-4}$	$< 10^{-4}$

4 DISCUSSION

4.1 Explanation of the Lack of Influence of Fill Density

From the results, it is evident that the fill density has no significant impact on the test outcomes. This unrealistic finding can be explained by delving into the print preparation process within the Eiger software, specifically by examining each individual material layer. Given that the samples were 2 mm thick, considering the specimens with carbon fibres with 0,125 mm layer height, each sample comprised 16 layers. The first and last 4 layers of each sample were identical, as they were printed with solid fill (Fig. 9) to ensure structural stability (a default setting in the software). Additionally, the software automatically incorporated 2 solid fill layers both before and after the fibre reinforced layer. Consequently, despite adjusting the fill density within the range of 28% to 54%, the majority of layers were printed with solid fill. It can be concluded that the adjustments in fill density were applied solely to the fibre reinforced layers, while the first and last 4 layers of each sample, as well as the layers before and after the fibre reinforced layer, were printed with solid fill for structural stability. Similar reasoning could be made also for specimens with glass and Kevlar fibres with 0,1 mm layer height. This limitation implies that the influence of fill density cannot be accurately assessed on samples of these dimensions with additional reinforced layers.

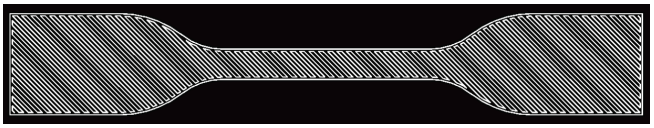


Figure 9 A solid fill layer example in a test specimen (four top and bottom layers, and two layers under and above a reinforced layer)

4.2 Influence of Different Fibres on Mechanical Properties

The results and obtained models clearly show that reinforcing the part with fibres leads to a significant improvement in the mechanical properties of the material. It can also be observed that different types of fibres offer varying degrees of mechanical enhancement, which can be somewhat related to their mechanical properties as presented in Tab. 2.

The results for the elastic modulus (Fig. 8c), where the highest values were obtained for carbon fibres, and similar values for glass and Kevlar fibres, can be directly associated with the tensile modulus values of the fibres (60 GPa for carbon, and 21 GPa and 27 GPa for glass and Kevlar, respectively). Additionally, observations for maximum strain (Fig. 8b), with the highest values for fibreglass, followed by Kevlar and carbon, clearly relate to the fibre values (3,8% for fibreglass, 2,7% for Kevlar, and 1,5% for carbon).

From the results for maximum stress (Fig. 8a), it can be observed that carbon and fibreglass provide the best and similar reinforcement, significantly higher than Kevlar. This finding is not entirely supported by the tensile strength values for the fibres, where carbon shows a value of 800 MPa, higher than the similar values of 610 MPa and 590 MPa for Kevlar and fibreglass, respectively. This indicates that the mechanical behaviour of fibre-reinforced composites is a result of complex interactions between the matrix and fibres. Consequently, it is not always possible to directly correlate the composite's mechanical properties to the individual properties of the fibre and matrix. Therefore, testing the final composite is the best practice to fully understand its performance. This is especially true when considering fracture behaviour, which is well known in composite micromechanics.

5 CONCLUSIONS

The study confirmed that ONYX 3D printing filament offers excellent mechanical properties compared to conventional materials used in additive manufacturing. Further reinforcement with fibres leads to additional enhancements in tensile strength, fracture deformation, and stiffness. Graphite and glass fibres provide the highest increase in strength, fibreglass offers the best enhancement in fracture deformation, and graphite in stiffness.

As the result of the research, a common linear model for three different fibre types (carbon, fibreglass and Kevlar) was established, correlating the degree of fibre reinforcement, characterized by the number of reinforced layers, with the tensile mechanical properties. While a unified model for different fibres type was advantageous in simplicity and practicality, individual models for each fibre type could have offered greater model agreement with experimental data. It was also intended to address the influence of the fill density in the investigation, however, due to how the internal structure of the part is crated in the used 3DP software, the most of the layers were printed with solid infill. Consequently, it was not possible to achieve the effects of fill density for the experimental set-up implemented. Addressing this issue would require larger specimens and different 3D structures in a part. Future studies could benefit from accounting for uncontrolled sources of variability through explicit blocking, which would strengthen the statistical validity of the results.

The proposed models for tensile strength, strain at maximum stress and modulus of elasticity could be very useful in design of 3D printed parts, offering simple solution for prediction of mechanical properties of a part. However, the models are limited to the parts with similar dimension and loading conditions to the conducted tensile tests. Also, model's accuracy is somewhat limited because of small number of tested specimens. Nevertheless, the study presented here could be used as a starting point and guide for further and more profound research.

6 REFERENCES

[1] Bhuvanesh Kumar, M., & Sathiya, P. (2021). Methods and Materials for Additive Manufacturing: A Critical Review on Advancements and Challenges. *Thin-Walled Structures, 159*, 107228. https://doi.org/10.1016/J.TWS.2020.107228

[2] Tofail, S. A. M., Koumoulos, E. P., Bandyopadhyay, A., Bose, S., O'Donoghue, L., & Charitidis, C. (2018). Additive Manufacturing: Scientific and Technological Challenges, Market Uptake and Opportunities. *Materials Today, 21*(1), 22–37. https://doi.org/10.1016/J.MATTOD.2017.07.001

[3] Park, S., Shou, W., Makatura, L., Matusik, W., & Fu, K. (Kelvin). (2022). 3D Printing of Polymer Composites: Materials, Processes, and Applications. *Matter, 5*(1), 43-76. https://doi.org/10.1016/J.MATT.2021.10.018

[4] Kristiawan, R. B., Imaduddin, F., Ariawan, D. U., & Arifin, Z. (2021). A Review on the Fused Deposition Modeling (FDM) 3D Printing: Filament Processing, Materials, and Printing Parameters. *Open Engineering, 11*(1), 639-649. https://doi.org/10.1515/eng-2021-0063

[5] Hsueh, M.-H., Lai, C.-J., Wang, S.-H., Zeng, Y.-S., Hsieh, C.-H., Pan, C.-Y., & Huang, W.-C. (2021). Effect of Printing Parameters on the Thermal and Mechanical Properties of 3D-Printed PLA and PETG, Using Fused Deposition Modeling. *Polymers, 13*, 1758. https://doi.org/10.3390/polym13111758

[6] Amza, C. G., Zapciu, A., Constantin, G., Baciu, F., & Vasile, M. I. (2021). Enhancing Mechanical Properties of Polymer 3D Printed Parts. *Polymers, 13*, 562. https://doi.org/10.3390/polym13040562

[7] Afonso, J. A., Alves, J. L., Caldas, G., Gouveia, B. P., Santana, L., & Belinha, J. (2021). Influence of 3D Printing Process Parameters on the Mechanical Properties and Mass of PLA Parts and Predictive Models. *Rapid Prototyping Journal, 27*(3), 487-495. https://doi.org/10.1108/RPJ-03-2020-0043

[8] Kačmarčik, J., Ćosić, A., Muminović, B., & Hadžalić, M. (2021). Ispitivanje zateznih svojstava uzoraka napravljenih 3D Printanjem FDM postupkom od PLA materijala. In *Proceedings of 12th Research/expert Conference with international participation (QUALITY 2021)*, Neum, Bosnia and Herzegovina, 16-19 June 2021,131-136. (in Bosnien)

[9] Sanei, S. H. R., & Popescu, D. (2020). 3D-Printed Carbon Fibre Reinforced Polymer Composites: A Systematic Review. *J. Compos. Sci., 4*(3), 98. https://doi.org/10.3390/jcs4030098

[10] Li, J., Durandet, Y., Huang, X., Sun, G., & Ruan, D. (2022). Additively Manufactured Fibre-Reinforced Composites: A Review of Me-chanical Behavior and Opportunities. *Journal of Materials Science & Technology, 119*, 219-244. https://doi.org/10.1016/j.jmst.2021.11.063

[11] Hetrick, D. R., Sanei, S. H. R., Bakis, C. E., & Ashour, O. (2021). Evaluating the Effect of Variable Fibre Content on Mechanical Properties of Additively Manufactured Continuous Carbon Fibre Composites. *Journal of Reinforced Plastics and Composites, 40*, 365-377. https://doi.org/10.1177/0731684420963217

[12] Sethu Ramalingam, P., Mayandi, K., Balasubramanian, V., Chandrasekar, K., Stalany, V. M., & Abdul Munaf, A. (2021). Effect of 3D Printing Process Parameters on the Impact Strength of Onyx – Glass Fibre Reinforced Composites. *Materials Today: Proceedings, 45*, 6154-6159. https://doi.org/10.1016/j.matpr.2020.10.467

[13] Petcharat, N., Wiangkham, A., Pichitkul, A., Tantrairatn, S., Aengchuan, P., Bureerat, S., Banpap, S., Khunthongplatprasert, P., & Ariyarit, A. (2023). The Multi-Objective Optimization of Material Properties of 3D Print Onyx/Carbon Fibre Composites via Surrogate Model. *Materials Today Communications, 37, 1*07362. https://doi.org/10.1016/j.mtcomm.2023.107362

[14] Piramanayagam, S. R., Kalimuthu, M., Nagarajan, R., Abdul Karim Sait, A. M., Krishnamoorthy, R. K., Ismail, S. O., Siengchin, S., Mohammad, F., & Al-Lohedan, H. A. (2021). Experimental Investigation and Statistical Analysis of Additively Manufactured Onyx-carbon Fibre Reinforced Composites. *J of Applied Polymer Sci, 138*, 50338, https://doi.org/10.1002/app.50338

[15] Radojičić, S., Konjatić, P., Katinić, M., & Kačmarčik, J. (2023). The Influence of Material Storage on Mechanical Properties and Deterioration of Composite Materials. *Tehnički vjesnik, 30*(5), 1645-1651. https://doi.org/10.17559/TV-20230308000422

[16] Standard ISO 527-1:2019: Plastics, Determination of Tensile Properties, Part 1: General Principles. International Organization for Standardization.

[17] Standard ISO 527-2:2012: Plastics, Determination of Tensile Properties, Part 2: Test conditions for moulding and extrusion plastics. International Organization for Standardization.

[18] Markforged, Mark Two. Available online: https://markforged.com/3d-printers/mark-two (accessed on 19.07.2024)

[19] Markforged, 3D Printing Materials. Available online: https://markforged.com/materials (accessed on 19.07.2024)

[20] Markforged, Material datasheet, Composites. Available online: https://www-objects.markforged.com/craft/materials/CompositesV5.2.pdf (accessed on 19.07.2024)

[21] Markforged, Eiger. Available online: https://markforged.com/software (accessed on 17.04.2024)

[22] Stat-Ease 360. Available online: https://www.statease.com/software/se360/ (accessed on 19.07.2024)

Authors' contacts:

Pejo Konjatić, PhD, Full Professor
Mechanical Engineering Faculty in Slavonski Brod, University of Slavonski Brod,
Trg Ivane Brlić Mažuranić 2, 35000 Slavonski Brod, Croatia
pkonjatic@unisb.hr

Josip Kačmarčik, PhD, Associate Professor
(Corresponding author)
Faculty of Mechanical Engineering in Zenica, University of Zenica,
Fakultetska 1, 72000 Zenica, Bosnia and Herzegovina
josip.kacmarcik@unze.ba

Marko Katinić, PhD, Associate Professor
Mechanical Engineering Faculty in Slavonski Brod, University of Slavonski Brod,
Trg Ivane Brlić Mažuranić 2, 35000 Slavonski Brod, Croatia
mkatinic@unisb.hr

Lamija Mešeljević, Assistant
Faculty of Mechanical Engineering in Zenica, University of Zenica,
Fakultetska 1, 72000 Zenica, Bosnia and Herzegovina
lamija.meseljevic@unze.ba

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240718091147

Subject review
Received: 2024-07-18, Accepted: 2025-01-20

Supervised Machine Learning Methods in Predicting English Premier League Game Outcomes

Mateo Vujčić, Tomislav Horvat*, Dejan Barešić, Dražen Crčić

Abstract: This paper explores the application of machine learning (ML) models to predict soccer game outcomes in the English Premier League. Utilizing data from fbref.com, various Features and target variables were analyzed to assess their influence on game results. The study implemented multiple ML algorithms, including logistic regression, Naive Bayes, decision trees, k-nearest neighbors, random forest, AdaBoost, and multilayer perceptron. The results highlighted significant variations in prediction accuracy across different teams and models, with the Random Forest model achieving the highest average accuracy. The findings underscore the importance of careful algorithm selection and data processing to enhance prediction precision. This research contributes to the field of sports analytics, providing insights that can be applied to improve tactical planning and decision-making in sports. Future work will focus on optimizing models and exploring additional data sources to further increase prediction accuracy.

Keywords: English Premier League; machine learning; predicting outcomes; soccer; supervised learning; Train&Test; validation methods

1 INTRODUCTION

Predicting outcomes in various fields is an increasingly essential and intriguing task. Whether it's making economic decisions, diagnosing health conditions, or devising sports strategies, the advantage provided by foreseeing future events is invaluable. Accurate predictions necessitate a thorough understanding and appropriate interpretation of relevant data, which can be structured, semi-structured, or unstructured. This introduction highlights the importance of outcome prediction and emphasizes the need for access to relevant data to enhance prediction accuracy.

The ability to predict outcomes is vital across diverse segments of life. In the business sector, predictions enable companies to make strategic decisions, adapt to market conditions, and minimize risks. In healthcare, accurate predictions contribute to timely diagnoses and personalized treatment plans, significantly improving patient care. In sports, data analytics is crucial for tactical planning and optimizing performance. This underscores the general significance of outcome prediction and the crucial role of relevant data in achieving precise predictions.

Access to relevant data is fundamental to achieving optimal results, regardless of the application area. Such data provides deeper insights into current conditions and forms the basis for informed decision-making. In business, relevant data includes information about consumer habits; in medicine, it encompasses biological and genetic markers; and in sports, performance analysis is key. Thus, success in outcome prediction relies on the efficient collection, analysis, and interpretation of relevant data.

The diversity of data presents both challenges and opportunities for proper collection, analysis, and interpretation. Understanding the different types of data is crucial for developing effective methods for outcome prediction. Structured data consists of organized information stored in defined structures like tables or databases, with clearly specified attributes and relationships. These data allow for easy searching, filtering, and analysis, forming the basis for quantitative and statistical analyses in business and science.

Semi-structured data, while partially organized, lack strictly defined attributes or relationships. Formats such as XML, JSON, and YAML represent these data, offering greater flexibility compared to strictly structured data but requiring more complex analysis methods. Unstructured data, on the other hand, includes information without clear organization, such as text documents, audio recordings, images, and videos. These data do not follow predefined patterns, necessitating the use of advanced processing techniques to extract relevant information [1].

The variety of data types offers valuable information but also presents processing challenges. For outcome prediction, combining different data types is often used for deeper insights and thorough analysis.

Predicting the outcomes of sports events has evolved into a significant area of research, especially in the realm of high-profile sports leagues like the EPL. In modern times, the vast amount of data collected from various sports events is used not only for detailed analysis but also for anticipating future game outcomes or even results.

After the introduction and a short motivation, section 2 provides an overview of existing papers from other researchers, focusing on studies related to the prediction of outcomes in sports, not exclusively limited to soccer. The obtained results, set of used features, validation methods will be analyzed and specific conclusions will be drawn. In Section 3 the input dataset and features of presented model will be described. Section 4 will outline the methodology, explaining the techniques and processes applied in the research. In Section 4, a predictive model will also be explained and presented. Section 5 presents and discusses the research results, while Section 6 concludes with a summary of the findings and suggestions for future research.

2 RELATED RESEARCH

Predicting outcomes in sports certainly represents a very interesting, but also challenging area. In this section, an overview of papers related to the outcome prediction in sports will be given, with an emphasis on soccer.

Tab. 1 provides an analysis of various studies conducted by researchers focusing on predicting sports outcomes [2-17]. These studies have primarily utilized supervised machine learning (ML) classification methods. However, the comparison of results is often challenging due to the use of different datasets across studies. The most commonly used validation method is data partitioning, with fewer researchers employing cross-validation. Due to the interconnected nature of sports events, cross-validation is not always suitable for sports predictions [18]. Best results are typically achieved using smaller datasets. This review analyzed 16 studies from the period 2008 to 2024, focusing on soccer. The most frequently studied league is the English Premier League, accounting for 55.56 % of the research. All studies presented in Tab. 1 utilize historical results, while some additionally incorporate data such as team rankings, weather conditions, individual player statistics, or even betting odds.

Table 1 Analyzed papers categorized by the publication year

#	Year	Input Dataset	Algorithms Used	Best Result (%)
[2]	2008	English Premier League (EPL) (2002-2007)	Neural Network	58.40
[3]	2011	Dutch Soccer Leagues (last 15 seasons)	Linear Regression	57.00
[4]	2011	Champions League	Neural Network, LogitBoost	68.80
[5]	2013	Spanish Soccer League (2008)	Bayesian Network	92.00
[6]	2014	EPL (2014)	Logistic Regression	93.00
[7]	2015	Dutch Eredivisie (2000-2012)	LogitBoost	56.10
[8]	2016	EPL (2010-2015)	Logistic Regression	69.50
[9]	2018	Spanish Soccer League (2012-2016)	Logistic Regression	71.63
[10]	2020	Multiple European Leagues (2013-2017)	Random Forest	75.62
[11]	2021	5 European Leagues (2006-2017)	Integrated Model	81.77
[12]	2021	UEFA Champions League (2016-2017)	Logistic Regression	81.00
[13]	2021	EPL (2005-2006)	Decision Tree, Neural Network	88.00
[14]	2022	Top 5 European Leagues (2014-2017)	Bayesian Network	92.01
[15]	2022	EPL (2018-2019)	SVM, Bayesian Network	61.32
[16]	2022	EPL (2019-2021)	Logistic Regression, k-NN, Random Forest	70.00
[17]	2024	World Cup (2022)	Random Forest	69.20

As can be seen in Tab. 1, the algorithms examined include logistic regression, Naive Bayes, decision trees, *k*-nearest neighbors (*k-NN*), random forest, LogitBoost, and multilayer perceptron (MLP) [19]. Each algorithm was tested on datasets comprising historical game results, team statistics, and factors related to home and away games. Most of these studies employ multiple ML algorithms, meaning the analyzed sample includes over 50 different prediction results. However, comparing these results is challenging or nearly impossible due to the use of different datasets and leagues of varying competitiveness. Only the best results are considered when different feature sets or datasets are used by the authors.

Fig. 1 shows a box plot of the range of results obtained by the analyzed studies in predicting soccer game outcomes. The box plot provides a detailed insight into the accuracy of different ML algorithms in predicting soccer outcomes. The results show significant variations in accuracy among the algorithms. Logistic regression is noted for its wide range of accuracy, indicating its adaptability to different datasets and processing methods. Naive Bayes shows consistent and stable results, while decision trees and k-NN have a broader accuracy range, suggesting their performance may depend on specific conditions and data features. Random forest often achieves high accuracy, making it a reliable choice for predicting soccer outcomes. LogitBoost also shows good results but does not reach the highest accuracy level of some other algorithms. Multilayer perceptron has consistent accuracy, indicating a potential need for further model fine-tuning or feature adjustment.

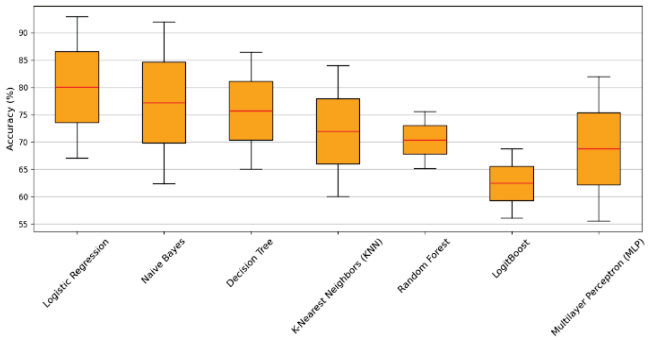


Figure 1 Box Plot Analysis of Results from Other Studies

These results underscore the importance of careful algorithm selection and data processing methods to achieve optimal outcomes in sports prediction. Using appropriate features and model adjustments is crucial for achieving high accuracy and reliability in predictions. Additionally, other researchers emphasize the significance of feature selection and feature extraction. Feature selection is often performed using specific methods and based on expert experience, aiming to identify the most relevant features that contribute to the predictive model performance. Feature extraction involves transforming raw data into informative and non-redundant features, which can enhance the predictive power. The analysis also shows that while some algorithms can achieve high accuracy under specific conditions, their overall success in practice is often determined by consistency and adaptability to different datasets. As noted, a large amount of data related to sports events is generated daily, both structured and unstructured. With the growing volume of data, the number of relevant databases containing various sports statistics also increases. Reviewing other researchers' work makes it clear that the main sources of information are mostly official sports organization websites. The study analyzed several studies focused on predicting sports outcomes or extracting useful information and patterns related to sports. Furthermore, as the availability and volume of sports data increase, so does the interest in using this data

for analytical purposes and predicting future outcomes. Sports organizations often provide detailed statistics and historical data through their official channels, enabling researchers to access valuable information for modeling and analysis. Many studies also use these data sources to develop and test various ML algorithms, striving to improve prediction accuracy in sports. There are also many other research studies, both scientific and review articles, not strictly related to predicting soccer outcomes, that provide valuable results and conclusions in predicting sports outcomes.

In addition to the mentioned papers concerning the outcome prediction in soccer, there is also a substantial body of review and scientific papers dedicated to predicting outcomes in various sports disciplines. These papers encompass a wide range of methodologies and approaches to enhance the accuracy and reliability of sports predictions. Furthermore, numerous scientific studies focus on the analysis of datasets specifically tailored for predicting sports outcomes, highlighting the importance of data quality and feature selection in developing robust predictive models. The authors in the study [20] and [21] provided an overview of the comparison of using ML algorithms and derived models in predicting outcomes in sports for various disciplines. The study [22] provides a comprehensive review of computer vision in sports. Studies [23-28] propose models for predicting outcomes in sports, specifically basketball and soccer. Study [29] discusses the long-term influence of technical and physical performance indicators, and situational variables on finale game outcome in soccer, while study [30] focuses on variations in the physical demands and technical performance of professional soccer teams.

Mentioned studies contribute to the growing knowledge base, offering valuable insights and advancements in sports analytics and outcome prediction.

3 DATA COLLECTION

The dataset was scraped from *fbref.com* platform to ensure it included all necessary information for an in-depth analysis of team performance, results, and key parameters throughout the season [31]. The *fbref.com* platform offers a robust database of soccer statistics, essential for evaluating various aspects of soccer games.

The features were chosen for their potential to influence game outcomes, providing valuable insights into different factors that could affect the game. The target variables represent the direct outcomes that the predictive models aim to predict, serving as a measure of the models' accuracy.

3.1 Dataset

The dataset used in this research includes features, presented in Tab. 2, and target variables, listed in Tab. 3, which are relevant to predicting soccer game outcomes. These features were selected for their potential influence on game results. Detailed data from English Premier League (EPL) games were carefully collected from the *fbref.com* platform, which provides extensive statistics on game outcomes, team performance, and individual player metrics. This database serves as a comprehensive foundation for the analysis conducted in this study. The selected features capture various factors that influence soccer game outcomes, while the target variables correspond to the results that the predictive models aim to predict.

In particular, the target variable in this analysis is the game outcome (win/loss/draw), which is derived from the combination of the features presented in Tab. 2. The model utilizes a variety of features, including but not limited to expected goals (*xg*), expected goals against (*xga*), possession percentage, and opponent team strength, to predict the final game result. The choice of features was motivated by their relevance and potential impact on the game outcome, allowing for a more robust and accurate prediction model.

Table 2 Features included in the dataset

Features	Description
venue_code	Encoded value representing the game venue (home or away)
opp_code	Encoded value representing the opposing team
hour	Hour of the day when the game was played
day_code	Encoded value representing the day of the week the game was played
xg	Expected goals for the team
xga	Expected goals against the team
poss	Possession percentage of the team

Table 3 Target variables included in the dataset

Target Variable	Description
gf	Goals for the team
ga	Goals against the team
sh	Total shots by the team
sot	Shots on target by the team

Expected goals (*xg*) is a metric used to assess the quality of scoring opportunities by considering factors such as shot location, angle, assist type, distance from goal, and defensive pressure. ML or regression models use historical data to estimate the likelihood of a shot resulting in a goal, assigning an *xg* value between 0 and 1, with higher values indicating greater goal-scoring probability. Similarly, expected goals against (*xga*) estimates the number of goals a team is likely to concede, based on the quality of shots taken by the opposing team. Like *xg*, *xga* incorporates factors such as shot location, type, defensive pressure, and goalkeeper positioning. The cumulative *xga* values reflect a team's defensive performance.

4 METHODOLOGY

ML has become a crucial tool for predicting sports outcomes, aiding coaches and managers in forecasting game results, assessing player or team performance, detecting injuries, and identifying emerging talents. Furthermore, such data is invaluable to the general public, especially in sports betting, offering a basis for more informed decisions [32].

The EPL, known for its high level of competitiveness, offers a wealth of data that is extremely valuable for analysts. Analyzing factors such as team formations, historical encounters, player injuries, and current form is essential for applying supervised ML methods to predict game outcomes [33]. Individual sports with binary outcomes, such as tennis or chess, are relatively easier to predict compared to soccer, which involves multiple variables and presents a much

greater challenge. However, studying unexpected sports events, like Leicester City's surprising Premier League title win in the 2015/2016 season, highlights the unpredictability of sports competitions and the potential for surprises on the field.

4.1 Prediction Tools

Scikit-learn is an open-source ML library for Python, known for its simple yet effective implementation of a wide range of ML algorithms and model evaluation tools. Due to its broad functionality in model creation, training, and evaluation, the library has gained popularity among researchers and practitioners in the field of ML [34].

Logistic regression, found in the *sklearn.linear_model* module, is used for binary and multi-class classification. It estimates the probability of class membership using a logistic function, transforming input features into probability estimates. This algorithm is particularly useful due to its simplicity and interpretability. Algorithms such as GaussianNB, MultinomialNB, and BernoulliNB from the *sklearn.naive_bayes* module are based on Bayes' theorem with an assumption of conditional independence between features. These algorithms are particularly effective for large datasets and are often used in classification problems such as spam filtering and sentiment analysis. The *sklearn.tree* module includes decision tree algorithms used for classification tasks. This algorithm splits data into subsets based on attribute values through a series of decision rules, resulting in easily interpretable structures. Despite their tendency to overfit, decision trees can be very effective when used with properly tuned parameters.

The *sklearn.neighbors* module's *k-NN* algorithm identifies the k-nearest neighbors to each data instance for classification tasks. This method is known for its simplicity and intuitiveness, making it highly effective for multi-class problems and complex data distributions [35]. Part of the *sklearn.ensemble* module, random forest is an ensemble method that creates multiple decision trees from randomly selected features and data samples. This method ensures diversification among the trees and reduces the risk of overfitting, making it widely used in fields such as biomedicine, ecology, and economics for classification and regression problems. The *sklearn.neural_network* module's *MLPClassifier* uses a structure of multiple layers of perceptrons for classification tasks. MLP employs supervised learning techniques and is trained using the backpropagation algorithm, making it effective for solving complex classification problems. These algorithms and tools from the *Scikit-learn* library facilitate the development and testing of ML models in various applications, enabling researchers and practitioners to focus on innovation and solving specific problems in the field.

4.2 Steps and Implementation

The process of developing a ML model using Python libraries, particularly *Scikit-learn*, involves several key steps. Block diagram of the proposed model is shown in Fig. 2. Through this methodology, the entire model development process is explained, from initial data processing to final model evaluation. To begin, required libraries for data manipulation and various ML models from *Scikit-learn* are imported. Game data is then loaded from a CSV file, setting the first column as the index. The date column is converted to datetime format, and a target variable target is created to indicate a win (1) or loss (0). Categorical variables venue and opponent are encoded, the hour is extracted from the time column, and the day of the week is derived from the date column.

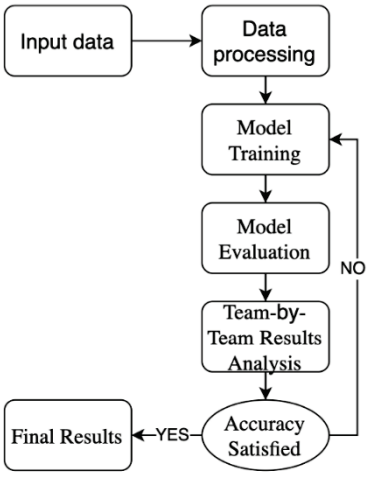


Figure 2 Block Diagram of the Proposed Prediction Model

Features for training the model are defined, and columns for calculating rolling averages are identified. The *rolling_averages* function sorts the data by date, computes rolling averages for specified columns, adds new columns with rolling averages, and removes rows without sufficient data. This function is applied to each team using the groupby and apply methods. The additional team index is then removed, and row indices in the *matches_rolling* data frame are reset.

Input dataset is split into training (1160 games before January 1, 2022) and testing sets (1132 games after that date). Various models for training are defined, including logistic regression, Naive Bayes, decision trees, k-NN, random forest, AdaBoost with logistic regression as the base model, and multilayer perceptron [19].

As already mentioned, several ML algorithms were used to predict match outcomes based on features outlined in Tab. 2. The target variable represents match results, where *W* indicated a win and *L* indicated a loss. Two approaches were used, Model-Level Accuracy and Team-Level Accuracy.

1) Model-Level Accuracy - The overall accuracy of each model was calculated as the ratio of correct predictions to total predictions across all teams.

$$Accuracy = \frac{N_{correct_predictions}}{N_{total_predictions}}. \quad (1)$$

2) Team-Level Accuracy - The accuracy for each team was calculated as the ratio of correct predictions to the total predictions for that specific team.

$$Accuracy = \frac{N_{gms_predicted} - N_{mssd_predictions}}{N_{gms_predicted}}. \quad (2)$$

The difference between goals scored (*gf*) and goals conceded (*ga*), although used for rolling averages and form indicators, did not directly affect the accuracy calculation, which was based on predicting match outcomes (win/loss). Results, including team-level accuracy for each model, are summarized in Tab. 4. The average accuracy of each model is also presented, showing the overall performance across teams.

A table to store results for each model and each team is created, with column names as model names and indices as unique teams from the test dataset. Each model is trained using training data, and outcomes for each team in the test dataset are predicted. Accuracy is calculated as the ratio of correctly predicted outcomes to the total number of games, and accuracy results are stored in the *team_results* table. Finally, average accuracy per team for each model is stored in the results table. The accuracy results of each model for each team and the average accuracy per model are printed using the print function.

Through this methodology, the application of ML techniques to real-world data is demonstrated. This approach enables the development of efficient models and a deeper understanding of the dynamics of sports competitions. The learned methods can be applied to various analytical challenges in many industries, opening possibilities for innovation and improving data-driven decision-making.

5 RESULTS

Fig. 3 highlight the complexity of predicting sports event outcomes, where factors such as team form, tactical decisions, and random events during the game can significantly impact prediction accuracy. The highest average accuracy per model is achieved by the Random Forest model at 77.01 %, indicating a satisfactory level of efficiency. However, the *k-NN* model shows the lowest average accuracy at 63.79 %, leaving room for improvement, especially in cases of teams with lower prediction accuracy.

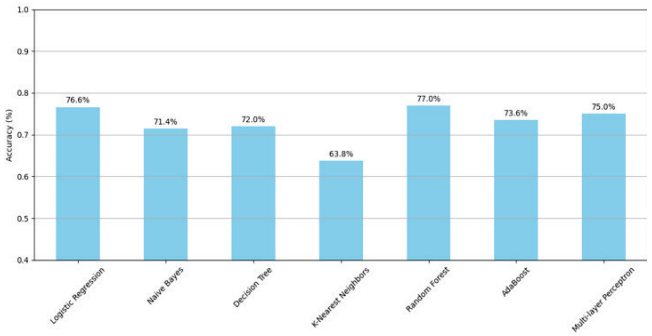


Figure 3 Algorithm Accuracy Comparison

Despite the best model achieving an accuracy of 77.01 %, a significant number of incorrect predictions were observed. This indicates the need for further analyses and potential improvements in the models to increase precision and reduce prediction errors. Improvements could include feature revision, parameter optimization, and testing alternative modeling approaches to achieve more reliable and accurate results.

Tab. 4 clearly illustrates the differences in accuracy of various ML models. The accuracy percentages above the bars allow for easy comparison of each model's effectiveness. An observed accuracy ranges from 47.06 % to 100.00 % suggests that there is room for improvement. Through further analysis of the results, it can be identified which models have the most potential for improvement and which optimization methods are most suitable for increasing overall accuracy and reducing prediction error.

Table 4 Detailed Accuracy of Algorithms

Team	*LR* (%)	*NB* (%)	*DT* (%)	*k-NN* (%)	*RF* (%)	*AB* (%)	*MLP* (%)
Arsenal	82.46	71.93	66.67	47.37	82.46	75.44	77.19
Aston Villa	70.69	63.79	72.41	53.45	72.41	68.97	70.69
Bournemouth	82.86	80.00	77.14	80.00	82.86	82.86	82.86
Brentford	67.24	63.79	68.97	68.97	62.07	68.97	63.79
Brighton	62.07	58.62	63.79	53.45	67.24	60.34	65.52
Burnley	86.36	72.73	68.18	68.18	86.36	90.91	86.36
Chelsea	64.29	50.00	67.86	66.07	64.29	66.07	64.29
Crystal P.	84.21	73.68	70.18	66.67	73.68	78.95	82.46
Everton	77.97	72.88	69.49	61.02	71.19	66.10	72.88
Fulham	65.79	63.16	73.68	68.42	78.95	65.79	68.42
Leeds United	75.86	81.03	81.03	67.24	75.86	81.03	68.97
Leicester City	74.14	72.41	72.41	68.97	72.41	70.69	74.14
Liverpool	77.19	70.18	66.67	59.65	82.46	68.42	75.44
Man. City	79.41	76.47	69.12	47.06	76.47	79.41	75.00
Man. United	74.14	63.79	70.69	51.72	77.59	67.24	74.14
Newcastle	73.68	59.65	61.40	54.39	71.93	63.16	73.68
Norwich City	89.47	84.21	78.95	73.68	84.21	84.21	84.21
Nottingham	74.29	77.14	80.00	71.43	85.71	74.29	71.43
Southampton	80.70	80.70	80.70	71.93	82.46	77.19	85.96
Tottenham	81.36	71.19	62.71	59.32	81.36	69.49	77.97
Watford	95.24	90.48	95.24	76.19	100.00	95.24	85.71
West Ham	77.19	77.19	70.18	61.40	71.93	71.93	77.19
Wolves	65.52	67.24	68.97	70.69	67.24	65.52	67.24

Additionally, the analysis of results from testing the ML model designed to predict EPL soccer outcomes is presented. Tab. 4 provides a detailed performance breakdown for each team, focusing on prediction accuracy. This visualization offers clear insights into the model's efficiency and identifies trends and anomalies in team performance. These results represent the average performance of the model for each club.

Analyzing the results presented in the table, significant variations in prediction accuracy between different teams are observed. Watford stands out with exceptionally high accuracy of 100.00 % for the Random Forest model, suggesting that the models successfully predicted their outcomes, possibly indicating certain predictability in their tactics or styles of play. Conversely, the lowest accuracy of 47.06 % was recorded for Manchester City using the *k-NN* model, indicating challenges in the models adapting to the variability characterizing their games.

6 CONCLUSION

The task of predicting outcomes across various domains has become increasingly important and captivating. From

making economic decisions and diagnosing medical conditions to formulating sports strategies, the ability to anticipate future events offers immense value. Achieving accurate predictions requires a deep understanding and proper analysis of relevant data.

In this study, the focus was on predicting soccer game outcomes in the EPL using various ML algorithms. The dataset was meticulously gathered from the *fbref.com* platform, encompassing a range of features and target variables relevant to soccer game outcomes. Features were chosen for their potential to influence game results, and the target variables were used to measure the models' accuracy.

The methodology section detailed the process of developing ML models using Python and the *Scikit-learn* library. The models tested included logistic regression, Naive Bayes, decision trees, *k-NN*, random forest, AdaBoost, and multilayer perceptron (*MLP*). Each model was evaluated based on its prediction accuracy for different teams in the league.

The results showed significant variations in prediction accuracy between different teams and models. The accuracy results ranged from 47.06 % (*k-NN* for Manchester City) to 100.00 % (Random Forest for Watford). Random Forest achieved the highest average accuracy at 77.01 %, demonstrating its effectiveness, while *k-NN* showed the most room for improvement with an average accuracy of 63.79 %.

These findings underscore the importance of careful algorithm selection and data processing methods to achieve optimal outcomes in sports prediction. Future research should aim to improve model accuracy by refining feature selection, optimizing hyperparameters, and exploring alternative modeling techniques. Additionally, integrating diverse types of data, such as real-time game statistics, player biometrics, and contextual factors like weather or crowd influence, could enhance prediction capabilities. Advanced processing techniques, including deep learning or hybrid models, may further boost performance and uncover hidden patterns in complex datasets. Such developments could lead to more reliable and insightful tools for decision-making in sports analytics [36].

In conclusion, this study demonstrates the potential of ML models to predict soccer game outcomes, highlighting the need for ongoing refinement and development. Accurate outcome prediction is vital across various fields, and the insights gained from this research can be applied to improve decision-making processes in sports and beyond. As data availability and computational methods continue to advance, the ability to predict outcomes with higher precision will become increasingly feasible, providing valuable benefits across multiple domains.

7 REFERENCES

[1] Spinner, J. (2018). *Web scraping with Python: Collecting more data from the modern web*. Sebastopol: O'Reilly Media.

[2] McCabe, A., & Travathan, J. (2008). Artificial Intelligence in Sports Prediction. In *Fifth International Conference on Information Technology: New Generations*. https://doi.org/10.1109/ITNG.2008.203

[3] Buursma, D. (2011). Predicting sports events from past results: Towards effective betting on football matches. In *14th Twente Student Conference on IT*.

[4] Hucaljuk, J., & Rakipović, A. (2011). Predicting football scores using machine learning techniques. In *Proceedings of the 34th International Convention MIPRO*.

[5] Owramipur, F., Eskandarian, P., & Mozneb Sadat, F. (2013). Football Result Prediction with Bayesian Network in Spanish League-Barcelona Team. *International Journal of Computer Theory and Engineering, 5*(5), 812-815. https://doi.org/10.7763/IJCTE.2013.V5.802

[6] Igiri, C., & Nwachukwu, E. (2014). An Improved Prediction System for Football a Match Result. *IOSR Journal of Engineering, 4*(12), 12-20. https://doi.org/10.9790/3021-04124012020

[7] Tax, N., & Joustra, Y. (2015). Predicting the Dutch Football Competition Using Public Data: A Machine Learning Approach. *Transactions on Knowledge and Data Engineering, 10*(10), 1-13.

[8] Prasetio, D., & Harlili, D. (2016). Predicting football match results with logistic regression. In *International Conference On Advanced Informatics: Concepts, Theory And Application (ICAICTA)*. https://doi.org/10.1109/ICAICTA.2016.7803111

[9] Zaveri, N., Shah, U., Tiwari, S., Shinde, P., & Teli, L. K. (2018). Prediction of Football Match Score and Decision Making Process. *International Journal on Recent and Innovation Trends in Computing and Communication, 6*(2), 162-165.

[10] Knoll, J., & Stübinger, J. (2020). Machine-learning-based statistical arbitrage football betting. *KI - Künstliche Intelligenz, 34*(1), 69-80. https://doi.org/10.1007/s13218-019-00610-4

[11] Stübinger, J., Mangold, B., & Knoll, J. (2020). Machine Learning in Football Betting: Prediction of Match Results Based on Player Characteristics. *Applied Sciences, 10*(1), 46. https://doi.org/10.3390/app10010046

[12] Ievoli, R., Palazzo, L., & Ragozini, G. (2021). On the use of passing network indicators to predict football outcomes. *Knowledge-Based Systems*. https://doi.org/10.1016/j.knosys.2021.106997

[13] Azeman, A. A., Mustapha, A., Razali, N., Nanthaamomphong, A., & Wahab, M. H. A. (2021). Prediction of football matches results: Decision forest against neural networks. In *The 8th International Conference on Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*. https://doi.org/10.1109/ECTI-CON51831.2021.9454789

[14] Razali, N., Mustapha, A., Mustapha, N., & Clemente, F. M. (2021). A Bayesian approach for major European football league match prediction. *International Journal of Nonlinear Analysis and Applications, 12*(Special Issue), 971-980.

[15] Rodrigues, F., & Pinto, Â. (2022). Prediction of football match results with machine learning. *Procedia Computer Science, 204*, 463-470. https://doi.org/10.1016/j.procs.2022.08.057

[16] Ren, Y., & Susnjak, T. (2022). Predicting Football Match Outcomes with eXplainable Machine Learning and the Kelly Index. In *School of Mathematical and Computational Sciences, Massey University, Auckland, New Zealand*.

[17] Demartino, R. M., Egidi, L., & Torelli, N. (2024). Alternative ranking measures to predict international football results.

[18] Horvat, T., Havaš, L. & Srpak, D. (2020). The Impact of Selecting a Validation Method in Machine Learning on Predicting Basketball Game Outcomes. *Symmetry, 12*(3), 431-446. https://doi.org/10.3390/sym12030431

[19] Kelleher, J. D., & Tierney, B. (2021). *Znanost o podacima*. Zagreb: Mate d.o.o. (in Croatian)

[20] Horvat, T., & Job, J. (2020). The use of machine learning in sport outcome prediction: A review. *Wiley interdisciplinary reviews, 1*, e1380. https://doi.org/10.1002/widm.1380

[21] Rico-González, M., Pino-Ortega, J., Méndez, A., Clemente, F. M., & Baca, A. (2023). Machine learning application in soccer: a systematic review. *Biology of Sport, 40*(1), 249-263. https://doi.org/10.5114/biolsport.2023.112970

[22] Naik, B. T., Hashmi, M. F., & Bokde, N. D. (2022). A Comprehensive Review of Computer Vision in Sports: Open Issues, Future Trends and Research Directions. *Applied Sciences, 12*(4429). https://doi.org/10.3390/app12094429

[23] Horvat, T., Job, J., Logožar, R., & Livada, Č. (2023). A Data-Driven Machine Learning Algorithm for Predicting the Outcomes of NBA Games. *Symmetry, 15*(4), 798. https://doi.org/10.3390/sym15040798

[24] Horvat, T., Job, J., & Medved, V. (2018). Prediction of Euroleague Games based on Supervised Classification Algorithm k-Nearest Neighbours. In Pezarat-Correia, P., Vilas-Boas, J., Rivera, O. et al. (Eds.), *Proceedings of the 6th International Congress on Sport Sciences Research and Technology Support* (pp. 203-207). Sevilla: SCITEPRESS. https://doi.org/10.5220/0006893502030207

[25] Horvat, T., & Job, J. (2019). Importance of the training dataset length in basketball game outcome prediction by using naive classification machine learning methods. *Elektrotehniški vestnik, 86*(4), 197-202.

[26] Bilek, G., & Ulas, E. (2019). Predicting match outcome according to the quality of opponent in the English premier league using situational variables and team performance indicators. *Journal of Performance Analysis in Sport, 19*, 930-941. https://doi.org/10.1080/24748668.2019.1684773

[27] Hsu, Y.-C. (2021). Using Convolutional Neural Network and Candlestick Representation to Predict Sports Match Outcomes. *Applied Sciences, 11*(6594). https://doi.org/10.3390/app11146594

[28] Zhang, Q., Zhang, X., Hu, H., Li, C., Lin, Y., & Ma, R. (2022). Sports Match Prediction Model for Training and Exercise Using Attention-Based LSTM Network. *Digital Communications and Networks, 8*, 508-515. https://doi.org/10.1016/j.dcan.2021.08.008

[29] Zhou, C., Calvo, A. L., Robertson, S., & Gómez, M.-Á. (2021). Long-Term Influence of Technical, Physical Performance Indicators and Situational Variables on Match Outcome in Male Professional Chinese Soccer. *Journal of Sports Sciences, 39*, 598-608. https://doi.org/10.1080/02640414.2020.1836793

[30] Akyildiz, Z., Nobari, H., González-Fernández, F. T., Praça, G. M., Sarmento, H., Guler, A. H., Saka, E. K., Clemente, F. M., & Figueiredo, A.J. (2022). Variations in the Physical Demands and Technical Performance of Professional Soccer Teams over Three Consecutive Seasons. *Scientific Reports, 12*, 2412. https://doi.org/10.1038/s41598-022-06365-7

[31] McDaniel, J. (2020). *Web Scraping with Python*. Birmingham: Packt Publishing Ltd.

[32] Alpaydin, E. (2021). *Strojno učenje*. Zagreb: Mate d.o.o. (in Croatian)

[33] Mohalder, R. N., Hossain, M. A., & Hossain, N. (2024). Classifying the Supervised Machine Learning and Comparing the Performances of the Algorithms. *International Journal of Advanced Research, 12*(1), 422-438. https://doi.org/10.21474/IJAR01/18138

[34] PMF. (2022). 6. Simpozij studenata doktorskih studija PMF-a. Zagreb. (in Croatian)

[35] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer. https://doi.org/10.1007/978-0-387-84858-7

[36] Herbinet, C. (2018). Predicting Football Results Using Machine Learning Techniques. London.

Authors' contacts:

Mateo Vujčić, bacc. ing. comp.
University North,
Trg dr. Žarka Dolinara 1, 48000 Koprivnica, Croatia
Armed Forces of Croatia,
Zagrebačka ul. 2, 43000 Bjelovar, Croatia
mateo7.vujcic@gmail.com

Tomislav Horvat, PhD, Assistant Professor
(Corresponding author)
University North,
104. brigade 3, 42000 Varaždin, Croatia
tomislav.horvat@unin.hr

Dejan Barešić, PhD, Assistant Professor
Croatian Military Academy "Dr. Franjo Tuđman",
Ilica 256b, 10000 Zagreb, Croatia
dejan.baresic@morh.hr

Dražen Crčić, mag. ing. el. eng.
University North,
104. brigade 3, 42000 Varaždin, Croatia
drazen.crcic@unin.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240810002519

Subject review
Received: 2024-08-10, Accepted: 2025-01-20

A Comprehensive Review of Artificial Intelligence Applications in Financial Market Predictions

Marcel Piskač

Abstract: This paper aims to provide a comparative summary of the possible applications of artificial intelligence (AI) in financial market movement forecasting. It also explores the historical development, key methodologies, and modern trends in AI-based market forecasting. Through the analysis of various studies and published research, this review aims to highlight the advantages, as well as the limitations of various AI techniques in the field. The findings provide actionable insight for researchers and practical users interested in using AI to analyse financial market trends.

Keywords: Artificial Intelligence (AI); financial markets; financial predictions; literature review; machine learning; neural networks

1 INTRODUCTION

The primary goal of this review is to provide a comprehensive overview of the various applications of artificial intelligence (AI) in financial market forecasting. By synthesizing existing research, the review aims to identify and compare key trends, methodologies and future directions in the field.

1.1 Background and Relevance

Forecasting financial market movements is a key challenge for all players in the field due to the inherent complexity and dynamic nature of markets. Traditional methods rely on economic theories and statistical models and are therefore limited [1] in their ability to capture non-linear relationships and complex patterns.

With the development of computational ability and the availability of extensive amounts of data, AI has come forward as a useful tool for analysing and predicting market movements. AI techniques such as machine learning and neural networks enable models to learn from large data sets, map extensive patterns, and make potentially more accurate predictions. Studies have shown that AI can outperform traditional methods in predicting stock price movements [2, 3], currency rates [4, 5] and other financial instruments [6].

The application of AI in finance is not limited to market forecasting. It is also used for risk management, fraud detection, and portfolio optimization. Understanding therefore the application of AI in financial markets has significant importance for researchers and practical users. For academic researchers, this review provides a comprehensive overview over the current trends, methodologies, challenges, and identifies areas for future research. For practical users in the field, it provides practical insight into successful AI applications, enabling them to make use of the latest techniques to improve analysis, and to make informed decisions.

Relevance is justified by the contribution to existing knowledge providing a comprehensive overview of the research up to date, and of the practical application in the financial sector. It also highlights key advantages as well as limitations of various AI techniques, essential for developing effective and reliable forecasting models.

1.2 Methodology

The methodology of this review is based on systematic research of relevant sources to collect and analyse studies related to the application of AI in the field. The search process involved several steps:

1) Selection of databases: the search was conducted through several academic databases, including IEEE Xplore, ScienceDirect, and Google Scholar. In addition, relevant field sources and reports were reviewed.
2) Keywords: Multiple combinations of keywords were used, including "AI", "financial market", "machine learning", "large language model", "LLM", "neural network", "forecasting", "pattern continuation", and "deep learning". The search was conducted exclusively in English.
3) Inclusion criteria: only studies published as of January 2015 were included to ensure that the data and conclusions were relevant. Studies had to be peer-reviewed and published in field-related journals or conferences. Also included are relevant field reports that provided conclusive insights.
4) Data analysis and synthesis: collected data was analysed and synthesized to identify repeating trends, methodologies, and results. The studies were further grouped according to topic, methodology, and the results of different studies are compared using a detailed descriptive method.
5) Validation and critical analysis: collected studies were critically analysed to assess their rigor, validity of result, and potential limitations. Special attention was paid to the identification of specific areas where further research is needed.

This methodology ensured a comprehensive and structured review of existing published work, so that the conclusions of this paper can be based on reliable and relevant sources.

2 OVERVIEW OF AI IN FINANCIAL MARKETS

The historical footprint of the application of AI in the field is traced from early statistical methods to modern machine and deep learning techniques:

- Early stages (1970s – 1990s): The first attempts are made in the application of computers in financial markets, which relied on statistical methods and existing econometric models [7]. These common forecasting methods were primarily regression analysis and time series; and the model used was primarily ARIMA [8]. In the 1980s first expert systems were developed, marking the beginning of proper AI application in the field. These systems [9] used rules which strictly followed expert knowledge to make objectively unbiased trading and portfolio management decisions.
- Development of neural networks (1990s – 2000s): Neural networks became popular in financial market applications due to their innate ability to learn from aggregated data and to recognize complex, extended-time-line patterns. Most notably, the use of multi-layer perceptrons (MLP) for stock price prediction. Such models had shown superiority [10] over traditional statistical methods in repeatable cases. In the early 2000s, further development of neural network training algorithms, most notably error backpropagation, measurably improved [11] the accuracy of the models applying them.
- The advent of machine learning (2000s - present): Algorithms such as support vector machines (SVM), random forests, and k-nearest neighbours (k-NN) came into utilization [12]. Simultaneously, an increase in the availability of financial data and the growth of computational power have enabled a more comprehensive utilization of these technologies. Today, deep learning and specific neural networks, most notably recurrent neural networks (RNN), and long-short-term memories (LSTM) have become the standard [13] in analysing large amounts of financial data, and therein predicting market movements.

2.1 Key AI Methodologies

Analysing the application of artificial intelligence in the field, different methodologies have been found [14, 15, 16, 17, 18] in use, each with its own measurable advantages and limitations. These included:

1) **Machine learning**

a) Supervised learning – uses labelled data to train models in predicting outcomes based on input data, so attempting to predict stock prices but also creditworthiness and the associated risk.

b) Unsupervised learning – analyses data without such predefined labels with the aim to discover hidden patterns and structures. Most notably it is used in customer segmentation, but in the field for the detection of trading anomalies.

c) Reinforcement learning – such models learn through making predetermined sequential decisions that aim to maximise long-term positive outcomes. It is mostly used in portfolio optimisation, as well as automated trading.

2) **Neural networks**

a) Multilayer perceptrons (MPL) – used for basic prediction and classification, through making used of one or more hidden layers in a simple neural network.

b) Recurrent neural networks (RNN) – aim to predict stock prices and financial indicators by working with time serious and sequential data.

c) Long-short-term memory (LSTM) – a type of RNN that is proven to be effective at capturing long-term dependencies in data, and therefore improving the accuracy of predictions.

3) **Deep learning**

a) Convolutional neural networks (CNN) – used to analyse and map patterns in a time series to a predetermined pattern.

b) Generative adversarial networks (GAN) – generate synthetic data to improve training and later testing of predefined trading strategies.

4) **Other noted techniques**

a) Genetic algorithms – utilised in the optimisation of trading strategies and portfolios, based on the natural selection principle.

b) Fuzzy logic – used to model uncertainty to define the conditions thereof, and later use these for decision-making.

c) Expert systems – these integrate existing expert knowledge and the resulting rules into automated trading decisions and supervise risk management.

These methodologies are a list of documented different approaches to using AI in the analysis and forecasting of financial market movements. By combining these, researchers and practical users can potentially develop robust models with ever more improved accuracy and efficiency of predictions and financial operations.

2.2 Current Trends and Applications

Applications of AI in the field continue to progress and become more prevalent with the advancement of technology itself and the ever-increasing availability of data. Current trends encompass several key areas and are transforming the way institutions as well as investors analyse their target market and make decisions.

The most prominently mentioned [14, 17, 18] trend is high frequency trading (HFT) which involves the execution of many transactions in very short time intervals (often milliseconds). AI algorithms are used to analyse real-time data and make near-instant trading decisions. Continually deep learning and machine learning algorithms analyse patterns and trends, enabling HFT systems to identify key events and anomalies in the market, to execute trades faster than a human trader. Consequently, this increases market liquidity and enables lower reaction times to market changes.

Sentiment analysis [19] is another prevalent application of AI in the field. It uses natural language processing (NLP) techniques to gauge the market sentiment, based on raw

textual data retrieved from several sources, including news articles and social media. Institutions use sentiment analysis not only to predict market movements, but also to identify potential opportunities and risks, and to make better informed decisions. Uniquely, NLP covers the human input into the field, thereby increasing accuracy of forecasts and the quality of decision-making by providing a deeper insight into sentiments, trends, and moods.

AI also shows significant impact on portfolio management. Using combinations of machine learning algorithms, investors attempt to optimise their portfolios by analysing historical data. These algorithms specifically identify optimal asset combinations that minimise risk and maximise returns. Such optimization algorithms are used to automatically balance and adjust portfolios close to real-time, all the time considering market conditions but also the individual goals which are defined in advance.

Automated trading [17, 18] is found as another key innovation driven by the development of AI. These systems use AI algorithms to analyse market data, making decision and executing transactions without human intervention. They are used in a variety of strategies, notably arbitrage, market making, and trend following. Such algorithms adjust their strategies in real-time, in accordance with predefined parameters, based on changing market conditions, thus increasing the efficiency of all future trading operations, eliminating human error (but not in the setting of parameters), and enabling 24/7 trading.

Separate from the trading aspect of the financial market, AI also seems to play a key role in risk management and fraud detection [20], where it shows significant potential. AI algorithms recognize patterns and anomalies which may indicate potential risks or fraudulent activity, therein providing financial institutions with the ability to better monitor real-time transactions, assess credit risk, and prevent fraud. Fraud detection systems specifically use machine learning to continuously learn and adapt to new threats, increasing the security of transactions, reducing losses due to fraud, and improving risk management.

2.2.1 Examples of Successful Implementations

Among the numerous examples of successful implementations of AI in the financial sector, few are notable and verifiable. These illustrate how AI can transform business processes and improve decision-making:

1) Renaissance Technologies [21], the hedge fund known for its heavy reliance on advanced mathematical models and AI, manages one of the most successful funds of this type, the Medallion Fund. Using machine learning, deep learning, as well as other AI techniques, they analyse large amounts of market data to identify patterns and so attempt to predict market movements. Their in-house model continuously adapts to new data, enabling the fund to achieve exceptional returns and minimise risks.
2) Two Sigma Investments [22], another hedge fund relying on AI and machine learning to make trading decisions, uses large data sets, including but not limited to market data, economic indicators, and alternative data sources such as weather and social media, to develop models that aim to predict market data. Their AI systems analyse this data in real-time, allowing them to react and to adjust quickly.
3) Bridgewater Associates [23], another from the group of hedge funds, uses AI in its flagship fund, the Pure Alpha Fund. Their AI algorithms aim to analyse macroeconomic data, market trends and historical patterns to make informed decisions. Their models help them identify long-term and short-term investment opportunities, but also to adapt to changes in the global economy and markets more time-efficiently.
4) Citadel [24], one of the global financial institutions that manage hedge funds but also market making operations, uses AI in its quantitative trading strategies. Their machine learning algorithms analyse large amount of market data to identify potentially profitable trading opportunities and optimise the end-users' trading strategies. Their models also adapt to changes in the market in real-time, allowing their quantitative trading strategies to generate consistent returns.
5) Goldman Sachs [25] implements AI through their Marquee trading platform, which is used to analyse market data for the purpose of making informed trading decisions. Marquee aims to predict the prices of stocks as well as other financial instruments, thus enabling individual traders to identify potentially profitable trading opportunities. Notably they also use AI for risk management and portfolio optimisation.

These selected examples of successful implementation of AI in well-established financial institutions show how the subject technologies can improve operational efficiency, and mitigate or reduce risks, while still catering personalised services to clients.

3 COMPARATIVE ANALYSIS OF AI TECHNIQUES

Convolutional Neural Networks (CNN) have shown notable performance in forecasting specifically stock market prices. They perform well in error analysis, and are more effective when combined with other models, most often paired with LSTM as a hybrid model. In a relevant study [26], the use of CNN combined with a variational autoencoder (VAE) for the purpose of feature extraction, showed superior performance in stock market predictions when compared to other methods.

Long-Short-Term Memory (LSTM) networks are widely recognized for their stability as well as accuracy in forecasting stock market returns. These outperform [27] traditional models like ARIMA and are most often integrated into ensemble models with similar neural networks. In the same study, one LSTM model was shown to produce lower errors compared to RNN and CNN models in forecasting NIFTY50 stock prices.

Integrating different AI methods into hybrid models has shown to be effective in improving prediction accuracy. One noted example [28] includes a combination of CNN, GRU (Gated Recurrent Unit) and LSTM, and has shown

significant improvement in stock market forecasting accuracy. Studied hybrid models in the study, such as the Random Forest + XG-Boost + LSTM ensemble, was particularly prominent for its effectiveness in predicting stock prices.

Explainable AI (XAI) methods like SHAP and LIME are currently being explored [29], with their purpose being providing transparency to the predictions of highly complex models such as LSTM and CNN. These seem to be the key to understanding the model's internal decision-making process, which is relevant for financial forecasting, as interpretability is as important as accuracy. [2]

Recurrent Neural Networks (RNN) are another noted time series forecasting model in the field. These have shown [30] to perform in various studies, very often with comparatively lower Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) value, when compared to models such as standalone LSTM.

Sentiment analysis, as a standalone technique, pulling current financial data from social media and news outlets, has proven [31] useful in improving the accuracy of forecasts. Market sentiment remains still a key driver of market movement. [5]

Ensemble models as their own category, models which combine multiple algorithms are noted for their robustness and superior performance in multiple financial forecasting variants. One such ensemble of neural networks has proven [32] to be superior to traditional models in predicting risk factors specifically in frontier markets.

Table 1 Comparative analysis of relevant findings

Technique	Accuracy	Stability	Interpretability	Notable indicators
CNN	High	High	Moderate	Superior feature extraction, effective in hybrid models
LSTM	Very high	Very high	Low	Lower errors when compared to RNN and CNN
Hybrid (CNN + LSTM)	Very high	Very high	Low to moderate	Significantly improved forecasting accuracy
XAI	High	N/A	High	Provides transparency in decision-making
RNN	High	High	Low	Lower RMSE and MAE values compared to others
Sentiment analysis integration	High	High	Low to moderate	Improved accuracy when combined with financial data
Ensemble models	Very high	Very high	Low	Outperforms traditional models in various scenarios

Labels "very high", "high", "moderate", and "low" used in the table were qualitative assessments based on the performed comparative analysis of the finding as reported in the reviewed studies. These were determined by examining key performance metrics, characteristics, and the overall conclusion presented in the research. These assessments were derived as follows:

1) **Accuracy**
a) Very high – consistently outperformed others regarding the forecast accuracy across multiple studies
b) High – showed strong performance in accuracy but were either less effective than "Very high" or had instances where the outperformance was not consistent
c) Moderate – performed adequately, with their accuracy less pronounced than in "Very high" or "High", and were generally outperformed by more advanced models
d) Low – had significantly higher error rates and lower accuracy compared to others

2) Stability
a) Very high – showed consistent performance across different datasets
b) High – performed reliably with most datasets but have shown slight variability
c) Moderate – had some degree of variability of performance across different datasets
d) Low – showed significant performance fluctuations, or were highly sensitive to changes in the input data

3) Interpretability
a) High – allowed clear insight into the decision-making process
b) Moderate – provided some level of interpretability
c) Low – had little to no ability to let users understand, or explain their decision-making process.

These qualitative labels are based on synthesizing information provided in the respective studies, where models were tested and compared. Performance indicators such as error rates, accuracy, and stability were the key factors in assigning these.

3.1 Advantages and Limitations

Among the main advantages of AI in the field, is the ability to parse large amounts of data in a short period of time, and to uncover complex patterns which may overwhelm traditional methods. CNNs excel at feature extraction, which allows for the recognition of complex patterns in large datasets. When combined with LSTMs most notably, they further improve on forecasting accuracy, as they successfully recognize spatial and temporal dependencies in the data.

LSTM networks are most effective at predicting a time series, which is the basis for all financial trend forecasting. Their reliability lies in the ability to retain tendencies over long sequences. On the other hand, the complexity of an LSTM can lead to overlearning [13], most commonly when small datasets are used.

When combined into a hybrid model, different AI techniques often achieve superior accuracy by combining the strengths of the different approaches. Consequently, the increased complexity and longer required training time can be a limitation. In hybrid models, alike LSTMs, their low interpretability introduces a further difficulty in working with them.

This problem of interpretability is met by XAI, created especially as its solution. However, in trying to bring insight into the decision-making process, to allow users to understand and justify the forecasts, it has the risk of oversimplifying the model.

Sentiment analysis aims to improve the predictive power of models, by enabling real-time insight into the market sentiment. In its application, it faces challenges such as messy and often unstructured data, difficulties in understanding nuances of natural language, and can ultimately fall victim to sarcasm or context-specific meaning, which result in misinterpretation.

Finally, ensemble models have shown a high degree of robustness and improved forecasting accuracy, in various financial scenarios. By averaging the predictions of several combined models, ensemble methods reduce risks and enhance the overall performance. This increased complexity however makes them more difficult to interpret and to implement, limiting their practical applicability.

4 CONCLUSION

The application of AI has fundamentally changed the approach to financial markets, offering tools that surpass traditional approaches in both accuracy and efficiency. Through the application of advanced techniques such as CNNs and LSTM networks, and ensemble models, financial institutions are now able to analyse extensive amounts of data with previously unattainable speed and measurably higher accuracy.

The comparative analysis presented here highlights significant leaps AI has made in the field, particularly in stock market predictions. CNNs have shown to be effective in feature extraction, making them valuable for identifying patterns within large datasets. LSTMs, which retain information over extensive data sequences, have become the staple for time series forecasting. Going even further, hybrid as well as ensemble models which integrate various AI techniques, combining the strengths of individual approaches, have an observable further enhanced predictive power.

Despite these advancements, the adoption of AI in the field remains challenging. The increasing complexity of models which is requires for greater accuracy brings the risk of overfitting and reduces the interpretability of the decision-making process ever more. These challenges require ongoing research and the refinement of AI techniques, which would ensure the provided predictions are not only accurate but also transparent and verifiable.

As AI continues to develop, its presence in the field is likely to expand further. Balancing this progress with the need for transparency will likely be crucial in realising its full benefits.

4.1 Advantages of AI in Financial Markets

Several key advantages were introduced into the field with the implementation of AI. Of the most significant benefits, is the ability to process and analyse large amount of data at unprecedented speeds. Models can efficiently parse large datasets, identify complex patterns and trends over long periods of time, which may be overlooked by a human analyst or a traditional statistical method.

In high-frequency trading (HFT) environments, where decisions are made in milliseconds, AI outperforms human traders by reacting to market movements at near instant times. This speed not only improved the efficiency of all trading operations, but at the same time also reduces the risk of missed opportunities due to delays in human cognitive decision making. The incorporation of real-time market sentiment analysis also provides a more comprehensive understanding of market dynamics.

In summary, the advantages are multifaceted, including enhanced data processing capabilities, forecasting accuracy, real-time decision-making, enhanced risk management, and the integration of sentiment analysis.

4.2 Limitations and Future Research

One of the primary challenges associated with AI is the required complexity of the models. Techniques such as LSTM and hybrid models often require extensive computational resources and time for training. At the same time, their complexity makes them difficult to interpret and manage. As models become more sophisticated, understanding how they arrive at specific decisions becomes vaguer. In the field, transparency and accountability are crucial, and the lack thereof hinders trust and adoption of AI driven decision-making.

Overfitting is a common problem with complex models (like LSTMs), where the models may perform exception on training data, but then fail to generalise to new, unseen data, particularly with smaller datasets. Mitigating overfitting required tuning and is very resource intensive. Sentiment analysis falls victim to unstructured and often ambiguous nature of textual data from social media and news outlets.

Future research for AI in the field should focus on several key areas: models need to become more interpretable. XAI addresses this issue but requires further advancement.

While hybrid and ensemble models show improved accuracy, their complexity and time required to train remain a hurdle. Research should aim to simplify these models without compromising on performance.

Improving sentiment analysis also seems to be essential. Its input into the decision-making process has a measurably positive effect but needs to be more reliable when it comes to interpreting the nuances of natural language.

Addressing the overfitting issue should be a priority. This could involve the introduction of new model validation techniques that ensure AI system generalise well to new data, particularly in the volatile financial markets.

By focusing on interpretability, simplification of complex models, refining sentiment analysis, and preventing overfitting, future research can help maximize the potential of AI in financial markets while minimizing its risks.

5 REFERENCES

[1] Tong, H. (2012). *Threshold models in non-linear time series analysis* (Vol. 21). Springer Science & Business Media.

[2] Weng, B., Lu, L., Wang, X., Megahed, F. M., & Martinez, W. (2018). Predicting short-term stock prices using ensemble

methods and online data sources. *Expert Systems with Applications, 112*, 258-273. https://doi.org/10.1016/j.eswa.2018.06.016

[3] Chopra, R., & Sharma, G. D. (2021). Application of artificial intelligence in stock market forecasting: a critique, review, and research agenda. *Journal of risk and financial management, 14*(11), 526. https://doi.org/10.3390/jrfm14110526

[4] Semaan, D. Y. (2014). *Forecasting exchange rates procedures: artificial intelligence or statistical techniques?* (Doctoral dissertation, Notre Dame University-Louaize).

[5] Zhang, Y., & Hamori, S. (2020). The predictability of the exchange rate when combining machine learning and fundamental models. *Journal of Risk and Financial Management, 13*(3), 48. https://doi.org/10.3390/jrfm13030048

[6] Tellez Gaytan, J. C., Ateeq, K., Rafiuddin, A., Alzoubi, H. M., Ghazal, T. M., Ahanger, T. A., ... & Viju, G. K. (2022). AI-Based Prediction of Capital Structure: Performance Comparison of ANN SVM and LR Models. *Computational intelligence and neuroscience*, *2022*(1), 8334927. https://doi.org/10.1155/2022/8334927

[7] Deboeck, G. J. (Ed.). (1994). *Trading on the edge: neural, genetic, and fuzzy systems for chaotic financial markets* (Vol. 39). John Wiley & Sons.

[8] Piccolo, D. (1990). A distance measure for classifying ARIMA models. *Journal of time series analysis*, *11*(2), 153-164. https://doi.org/10.1111/j.1467-9892.1990.tb00048.x

[9] Zenios, S. A. (1999). High-performance computing in finance: The last 10 years and the next. *Parallel Computing*, *25*(13-14), 2149-2175. https://doi.org/10.1016/S0167-8191(99)00083-6

[10] Devadoss, A. V., & Ligori, T. A. A. (2013). Forecasting of stock prices using multi layer perceptron. *International journal of computing algorithm*, *2*(1), 440-449.

[11] Hayward, S. (2006). Financial modeling and forecasting with an evolutionary artificial neural network. In *Computational Economics: A Perspective from Computational Intelligence* (pp. 1-33). IGI Global. https://doi.org/10.4018/978-1-59140-649-5.ch001

[12] Bansal, M., Goyal, A., & Choudhary, A. (2022). A comparative analysis of K-nearest neighbor, genetic, support vector machine, decision tree, and long short term memory algorithms in machine learning. *Decision Analytics Journal*, *3*, 100071. https://doi.org/10.1016/j.dajour.2022.100071

[13] Sherstinsky, A. (2020). Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D: Nonlinear Phenomena*, *404*, 132306. https://doi.org/10.1016/j.physd.2019.132306

[14] Chopra, R., & Sharma, G. D. (2021). Application of artificial intelligence in stock market forecasting: a critique, review, and research agenda. *Journal of risk and financial management*, *14*(11), 526. https://doi.org/10.3390/jrfm14110526

[15] Nair, B. B., & Mohandas, V. P. (2015). Artificial intelligence applications in financial forecasting – a survey and some empirical results. *Intelligent Decision Technologies*, *9*(2), 99-140. https://doi.org/10.3233/IDT-140211

[16] Zavadskaya, A. (2017). Artificial intelligence in finance: Forecasting stock market returns using artificial neural networks. *Master's thesis*. Hanken School of Economics.

[17] Alex Avelar, E., & Jordão, R. V. D. (2024). The role of artificial intelligence in the decision-making process: a study on the financial analysis and movement forecasting of the world's largest stock exchanges. *Management Decision*. https://doi.org/10.1108/MD-09-2023-1625

[18] Abdollahi, J., & Mahmoudi, L. (2021). Investigation of artificial intelligence in stock market prediction studies. In *10th International Conference on Innovation and Research in Engineering Science*.

[19] Mishev, K., Gjorgjevikj, A., Vodenska, I., Chitkushev, L. T., & Trajanov, D. (2020). Evaluation of sentiment analysis in finance: from lexicons to transformers. *IEEE Access*, *8*, 131662-131682. https://doi.org/10.1109/ACCESS.2020.3009626

[20] Hassan, M., Aziz, L. A. R., & Andriansyah, Y. (2023). The role artificial intelligence in modern banking: an exploration of AI-driven approaches for enhanced fraud prevention, risk management, and regulatory compliance. *Reviews of Contemporary Business Analytics*, *6*(1), 110-132.

[21] Zuckerman, G. (2019). *The man who solved the market: How Jim Simons launched the quant revolution*. Penguin.

[22] Mackintosh, J. (2017). *The Quants Run Wall Street Now*. The Wall Street Journal. https://www.wsj.com

[23] Brown, A., & Farrell, M. (2021). *The World's Largest Hedge Fund Is Building a "Machine Army" at Bridgewater*. Bloomberg. https://www.bloomberg.com

[24] Borzykowski, B. (2019). *How hedge fund Citadel gets an edge in recruiting and retaining top talent*. CNBC. https://www.cnbc.com

[25] Tracy, R. (2018). *Goldman Sachs Wants to Build a Netflix for Data*. The Wall Street Journal. https://www.wsj.com

[26] Parida, A. K., Kumari, R., Singh, A. R., Bandyopadhyay, A., & Swain, S. (2023, December). Stock market prediction under a deep learning approach using Variational Autoencoder, and kernel extreme learning machine. In *IEEE OITS International Conference on Information Technology (OCIT2023)*, 156-161. https://doi.org/10.1109/OCIT59427.2023.10430718

[27] Fathali, Z., Kodia, Z., & Ben Said, L. (2022). Stock market prediction of Nifty 50 index applying machine learning techniques. *Applied Artificial Intelligence*, *36*(1), 2111134. https://doi.org/10.1080/08839514.2022.2111134

[28] Sonkavde, G., Dharrao, D. S., Bongale, A. M., Deokate, S. T., Doreswamy, D., & Bhat, S. K. (2023). Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications. *International Journal of Financial Studies*, *11*(3), 94. https://doi.org/10.3390/ijfs11030094

[29] Valensky, D. & Mohaghegh, M. (2023, December). Evaluating Transparency: A Cross-Model Exploration of Explainable AI in Financial Forecasting. In *IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE2023)*, 1-6. https://doi.org/10.1109/CSDE59766.2023.10487732

[30] Sharma, P., Sharma, C., & Mathur, P. (2023, August). Machine Learning-based Stock Market Forecasting using Recurrent Neural Network. In *9th IEEE International Conference on Smart Computing and Communications (ICSCC2023)*, 600-605. https://doi.org/10.1109/ICSCC59169.2023.10335083

[31] Al Ridhawi, M. & Al Osman, H. (2023). Stock Market Prediction from Sentiment and Financial Stock Data Using Machine Learning. In *Canadian AI*. https://doi.org/10.21428/594757db.40c1a462

[32] Akter, M. S., Shahriar, H., Chowdhury, R., & Mahdy, M. R. C. (2022). Forecasting the risk factor of frontier markets: A novel stacking ensemble of neural network approach. *Future Internet*, *14*(9), 252. https://doi.org/10.3390/fi14090252

Authors' contacts:

Marcel Piskač
DOBA Business School,
Prešernova 1, 2000 Maribor, Slovenia
marcel.piskac@gmail.com

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240831112713

Subject review
Received: 2024-08-31, Accepted: 2025-04-28

Realization Lines with Inkjet Head KM1024iLHE-30 and Additional Applying Metalized Foil

Igor Majnarić, Marko Morić*, Marko Kadija, Krunoslav Hajdek

Abstract: The Inkjet printing technology has found widespread application across diverse industries, particularly in graphic and textile printing. The crucial role of the Inkjet head in achieving precise prints cannot be overstated. The selection of the inkjet head significantly influences the homogeneity of the printed surface and the realization of screen elements. Furthermore, the inkjet printing principle's potential for adhesion of metalized foils through a digital foiling process necessitates a detailed investigation into the influence of commercial piezo Inkjet heads. In this study, we present a method for testing the efficiency of achieving four-sized metalized gold foil printing lines on standard mat-coated cardboard printing substrates, both in the machine direction and cross direction. The experiment involves the gradual variation of the amount of UV-cured Inkjet varnish. Utilizing three different thicknesses of UV-cured varnish (21 µm, 36 µm, and 51 µm) in conjunction with modular piezo inkjet heads KM1024 iLHE-30 and print speeds according to varnish thickness (0,300 m/s, 0,24 m/s and 0,182 m/s). Research demonstrates that a 21 µm layer thickness of UV varnish combined with cross-direction movement ensures optimal production of digital metalized foil lines. Additionally, our results confirm the superior potential of a larger varnish application for line reproduction and enhanced contrast.

Keywords: digital foil lamination; line quality; piezo inkjet head; UV cured varnish

1 INTRODUCTION

Inkjet technology has a wide range of applications in printing various media products. These machines use non-impact printing technology, where liquid ink or dye droplets are directly distributed onto the printing surface. The dominant technology among inkjet printers is Piezo Inkjet, which generates droplets on demand through the mechanical displacement of a dynamic piezoceramic element causing pressure and the ejection of ink droplets.

Today, piezo Inkjet machines are used in office and business printing, commercial printing, decorative printing, packaging printing, printing of electronic components, and printing on textile materials. In the field of decorative printing, there has been a significant increase in the use of Inkjet printing for bonding metalized foils through digital foil transfer technology.

Additionally, Inkjet technology has the potential to be used in various non-traditional printing applications such as film formation, 3D modeling, painting, functional material coating, granulation, biotechnology, and medicine. To ensure quality printing on various substrates, the distribution of the liquid ink must be strictly controlled. Therefore, there are many commercial piezo inkjet heads available on the market, designed to generate the exact volume of a liquid drop in a time interval of a few microseconds.

2 THEORETICAL PART

During a visit to DRUPA 2024, it was observed that the most modern Inkjet machines are dominated by four global manufacturers of high-quality piezo Inkjet heads. These leading manufacturers and series include Kyocera (KJ4A and KJ4B series), Ricoh (TH and MH series), FujiFilm (Skyfire, Samba, StarFire, Sapphire, Polaris, Emerald, Galaxy, and Nova series), and Konica Minolta (Shear 3 cycles, Shear independent and Bend models).

2.1 Kyocera Inkjet Heads

The KJ4A series (models KJ4A-0300, KJ4A-RH, KJ4A-AA, KJ4A-TA) is designed for ink with relatively high viscosity, like UV curable ink. This series is best suited for various print media applications, such as overprint, labels, and others. Two resolutions are available (300 × 600 dpi and 600 × 600 dpi) with a jetting frequency of 20 to 30 kHz and the size of dynamically changing droplets in the volume range from 3 pl to 17 pl. This provides two print speeds of 50,8 m/min and 76,2 m/min.

The KJ4B series (models KJ4B-EX600-RC, KJ4B-EX1200-RC, KJ4B-EX600, KJ4B-1200, KJ4B-0300, KJ4B-YH, KJ4B-QA) is designed for high-speed printing with water-based ink. This series delivers excellent productivity not only in single-pass applications such as transaction printing, document printing, and security printing using MICR ink, but also in multi-pass applications, including textile printing. Three resolutions are available in this series: 300 × 600 dpi, 600 × 600 dpi, and 1200 × 1200 dpi. This means operating with a 30, 40, 64 kHz jetting frequency and the size of dynamically changing droplets in the volume range from 2 pl to 16 pl. Two printing speeds are possible: 76,2 m/min and 101,6 m/min. Machines can implement water-based ink with stable recirculation technology by choosing an Inkjet head printing option. [6]

2.2 Ricoh Inkjet Heads

The Japanese company Ricoh is one of the world's largest manufacturers of piezo inkjet heads, which have found applications in all sectors of printing and coating. To make this possible, there are three factory designs of Ricoh Inkjet heads that achieve droplet volumes ranging from 2,5 pl to 90 pl: TH series (2 models), MH series (9 models), and GH series (one model).

2.2.1 Ricoh TH Series

This series of MEMS Inkjet heads consists of the following models: Ricoh TH5241 and Ricoh TH6310F.

Their common feature is a thin-film piezoelectric transducer with bend mode, possibly using all types of Inkjet inks. Therefore, they are intended for different applications. One Ricoh TH5241 head contains 1280 nozzles arranged in a 4 x 320 configuration. This configuration enables printing in multi-drop technology with droplet sizes from 3 to 12 pl and three configurations: 4 inks at once in a resolution of 300 dpi (for each CMYK color one line), spot colors at a resolution of 600 dpi (2 lines for each ink) and one color at a resolution of 600 dpi (4 lines for one color). The model TH6310F features 600 dpi printing with two heads joined together in one standard housing (1600 nozzles = 8 rows × 100 nozzles × 2 heads). This configuration enables a greater printing distance compared to the printing surface, which is used for printing on materials such as textiles, labels, and corrugated cardboard. The droplet volumes of 3 pl to 12 pl and 7 pl to 35 pl achieved by these models are significant as they determine the resolution and quality of the printed material. Ink jetting is 80 kHz at 5 pl binary driving or 40 kHz at 5 pl, 10 pl 18 pl multi-drop driving. It also achieves the high productivity of 100 m/min print speed at 600 dpi.

2.2.2 Ricoh MH Series

The MH series Inkjet heads operate on the principle of piston pulser with metallic diaphragm plate and are also known on the market as Gen 4 generation (Ricoh MH2420, Ricoh MH2620/2820, Ricoh MH2810F/2910F) and Gen 5 generation (Ricoh MH5220, Ricoh MH5320/5340, Ricoh MH5420/5440, Ricoh MH5421/5441, Ricoh MH5422/5442 and Ricoh MH5421F/5421MF).

The heads of the Gen 4 generation are characterized by variable drop volume droplet formation technology (the possibility of connecting three smaller drops into one larger one) with the durability of nozzles arranged in only two rows (2 x 192 nozzles in a line). This enables the heads to apply only one or two Inkjet inks with a 30 and 20 kHz jetting frequency. The Ricoh MH2420 model thus has a heater that regulates the color temperature up to 60° C, creating volume drops from 7 pl to 35 pl. The Ricoh MH2620/2820 model has a more robust ink heater (from 80° C), which enables the application of various Inkjet inks to generate larger droplets from 15 pl to 45 pl or 27 pl to 54 pl. Model MH2810F/2910F is a stainless steel print head that realizes the most significant Ricoh volume of released droplets from 27 pl to 50 pl and from 50 pl to 90 pl. With this characteristic (with appropriate heating), the possibility of applying highly pigmented Inkjet inks, which ensure the formation of a more extensive inks layer, has been realized.

Ricoh heads of the Gen 5 generation are characterized by their installation in Inkjet machines of high production and high reproduction quality. This means they can operate in binary droplet formation mode (volumes from 2,5 pl to 7 pl) and grayscale mode (volumes from 7 pl to 35 pl). The heads of this series have a native print resolution of 600 dpi and contain 1280 nozzles (arranged in a configuration of 4 x 150 nozzles).

With such a configuration, it is possible to print four inks at once (one line per color = 150 dpi), two spot inks (two lines per color = 300 dpi), and one ink (four lines per color = 600 dpi). The Ricoh MH5220 model is intended only for the use of UV-curable inks. It prints a maximum of two colors during operation while generating the smallest droplet volumes from 2,5 pl to 9 pl. The Ricoh MH5320/5340 model has a stainless steel housing with built-in heaters, which makes it capable of working with most common Inkjet inks. With a high frequency of jetting (50 kHz) and a small volume of print dripping (from 5 pl to 15 pl), the head is suitable for easy connection in more comprehensive modules, enabling a color print of larger dimensions. The Ricoh MH5420/5440 model is used with all Inkjet inks and does not have a pronounced drip rate (max. 30 kHz). However, the head can generate high-quality images thanks to the grayscale mode that creates variable droplet sizes from 7 pl to 35 pl. The Ricoh MH5421/5441, Ricoh MH5422/5442, and Ricoh MH5421F/5421MF models have almost identical performance (volume 7 pl, 14 pl, and 21 pl) and are used in the most productive Inkjet machines. A slight difference is that water-based inks are applied (Ricoh MH5421/5441). It can recycle all types of Inkjet inks with a viscosity range of 6 pl to 11 mPa·s (MH5421F/5421MF) with greater precision in connecting to larger modules and making prints with enhanced non -wet-coating (Ricoh MH5422/5442).

2.2.3 Ricoh GH Series

The Ricoh GH series, represented by the Ricoh GH2220 head, operates on the piston pulser with a metallic diaphragm plate principle. This series is known for its excellent cost performance, making it a valuable asset for the sign graphic textile, coding, and marketing markets. The GH2220 head contains 384 nozzles that form droplets with a volume of 3 pl to 5 pl (binary mode) or 21±2 pl (grayscale mode) at a frequency of jetting 30 kHz. While the head does not have heaters, it is suitable for using water and solvent-based inks, making it a cost effective choice for various applications.

2.3 FUJIFILM Dimatix Inkjet Heads

Formerly known as Spectra Corporation, now operating as FUJIFILM Dimatix, Inc., we are a leading global supplier of drop-on-demand Inkjet printheads for commercial and industrial applications. With a robust portfolio of 37 types of piezo Inkjet heads, we stand out with our unique features and capabilities. Fujifilm printheads, divided into nine commercial series: Skyfire, Samba, Starfire, Sapphire, Polaris, Emerald, Galaxy, Nova, and S series, are a testament to our innovation and commitment to quality.

2.3.1 FujiFilm Skyfire Series

The SKYFIRE SF600 head currently represents this series with a native resolution of 600 dpi (1536 addressable nozzles jets in one module), which can form droplets from 5 pl to 25 pl (in binary and grayscale mode). Applying multi-pulse jetting technology with a frequency of 50 kHz, it is possible to achieve four printing speeds (66, 110, 130, and 264 m/min).

This speed directly corresponds with varying the printing resolution (600 × 1200 dpi, 600 × 720 dpi, 600 × 600 dpi, and 600 × 1200 dpi). This directly affects the ink laydown: 5,4 g/m^2, 10,8 g/m^2, 13 g/m^2, 21,7 g/m^2. As the head with Silicon

MEMs w/protective REDIJET coating that applies an additional heater with a thermistor, it is possible to print Aqueous, Solvent, and UV-curable inks. [8]

2.3.2 FujiFilm Samba Series

In the current fastest Inkjet machines, we usually find FujiFilm heads of the Samba G3L and Samba G5L. It is about silicon MEMs W heads with a native resolution of 1200 dpi (containing 2048 nozzles). Applying multi-pulse jetting technology with a frequency higher than 100 kHz, it is possible to achieve three printing speeds: 461 m/min (with a resolution of 1200 × 300 dpi), 230 m/min (with a resolution of 1200 × 600 dpi) and 115 m/min (with a resolution of 1200 × 1200 dpi). Applying binary and grayscale mode, Samba models achieve different volumes of formed droplets (Samba G3L from 2,4 to 10 pl; Samba G5L from 3,5 to 15 pl), ultimately affecting the achieved ink laydown. Thus, the more precise Samba G3L head will give smaller ink layers on paper substrates: 1 g/m^2 (resolution 1200 × 300 dpi), 2 g/m^2 (resolution 1200 × 600 dpi) and 4 g/m^2 (resolution 1200 × 300 dpi). Because of this, Samba heads can apply demanding latex Inkjet inks in addition to standard Aqueous, Solvent, and UV UV-curable inks (only Samba G3L heads).[9]

2.3.3 FujiFilm StarFire Series

The FujiFilm StarFire series is a testament to adaptability, offering the most significant number of construction variations. This characteristic has produced 23 Starfire models by application, divided into three groups (series A, series C, and series F). All StarFire heads of the SG 1024 series contain a metal plate with nozzles; a total number of 1024 print nozzles (8 lines × 128 nozzles) achieve a native print resolution of 400 dpi or 200 dpi (if the same head applies two inks). The heads of the SG 600 series achieve a higher output resolution (600 dpi or 300 dpi if two inks are printed), which was completed by installing a metal plate with 1536 nozzles (12 print lines × 128 nozzles). All models use binary and grayscale image generation methods and contain a system for dual ink recycling. This requires a viscosity of Inkjet inks in the 8 to 20 mPa·s range.

The heads of the A series are characterized by different dripping frequencies (50 kHz, 30 kHz, 20 kHz), which also achieve different droplet sizes. Thus, Starfire A series models will achieve the following native droplet volumes: SG1024/XSA (6 pl), SG1024/SA (12 pl), SG1024/MA (30 pl), SG1024/LA (80 pl), SG1024/SA-2C (12 pl), SG1024/MA-2C (30 pl), SG1024/LA2Ci (80 pl), SG600/XSA (6 pl) SG600/SA (12 pl) and SG600/SA2C (12 pl). This ensures excellent application of most UV curable, organic solvents and aqueous inks.

The heads of the C series are also characterized by differently realized ink droplet sizes and their native dripping frequencies of 50 kHz, 30 kHz, and 20 kHz. Thus, Starfire C series models will achieve the following native droplet volumes: SG1024/XSC (6 pl), SG1024/SC (12 pl), SG1024/MC (30 pl), SG1024/LC (80 pl), SG1024/MCHF (30 pl), SG1024/LCHF (80 pl), SG1024/AQMH (30 pl), SG1024/AQMH (80 pl), SG600/XSC (6 pl), SG600/SC (12 pl) and SG600/MC2C (30 pl). This ensures an excellent application of oil and solvent-based ceramic inks.

Our f series heads are designed for unique applications in sand casting and metal sintering. These heads are specifically engineered to apply aggressive fluids, preventing clogging of nozzles and ensuring a more extensive ink layer (laydown 50 g/m^2). To achieve this, our StarFire F series models produce larger droplets, with native volumes of 30 pl (SG1024/M3F head) and 80 pl (SG1024/L3F head).

2.3.4 FujiFilm Polaris Series

The Fuji Inkjet head series consists of three models: Polaris PQ-512/85 AAA, Polaris PQ-512/35 AAA, and Polaris PQ-512/15 AAA. All three Printheads have 512 independent nozzles arranged in multiple rows at 200 dots-per-inch spacing for single-color operation or 100 dots-per-inch spacing for two-color operation. The printhead is designed for binary jetting using an 80 pl drop size (PQ-512/85 AAA), 35 pl (Polaris PQ-512/35 AAA), and 15 pl native drop size (PQ-512/15 AAA). The nominal drop velocity for jetting fluids is 8 m/s thanks to two electrical connectors, which contain 8 serial data lines that support fire frequencies over 20 kHz, 30 kHz, and 40 kHz. Using VersaDrop jetting technology, the printhead is capable of adjustable binary operation at drop sizes up to 150 pl with a frequency of 10 kHz (PQ-512/85 AAA), 60 pl with a frequency of 18 kHz (Polaris PQ-512/35 AAA), and 30 pl with the frequency of 25 kHz (PQ-512/15 AAA). Polaris Printhead is compatible with a broad range of Inkjet inks (8 mPa·s to 20 mPa·s), including UV-curable, aggressive organic solvent and aqueous inks. [11]

2.3.5 FujiFilm Sapphire Series

Unlike the previously mentioned Inkjet heads, the MEMS Sapphire series has a slightly lower printing resolution (100 dpi), i.e., the number of addressable jets nozzles per module is 256. The series consists of three models: Sapphire QS-256/80 AAA, Sapphire QS-256/30 AAA, and Sapphire QS-256-10 AAA. The printheads are designed for grayscale jetting using a 10 pl, 14 pl, 30 pl, or 80 pl native drop size. As with the Polaris series, the nominal drop velocity for jetting inks is achieved with 8 m/s, which enables ink viscosity from 8 mPa·s to 20 mPa·s. Printhead operational temperature can be up to 90 °C. This allows for applying a broad range of printing inks, including UV curable and aqueous inks and associated maintenance fluids. [12]

2.3.6 FujiFilm Emerald Series

The Emerald QE family of printheads uses field-proven materials and a robust design to deliver a long service life. Although the printing characteristics are similar to the Sapphire series (printing resolution 100 dpi and number of addressable jets per module 256), the difference is evident in the metal nozzle plate technology. The series consists of two models: Emerald QE-256/30 AAA and Emerald QE-256/80 AAA. Their difference is in the precision of the formed droplet, which in the QE-256/30 AAA model is 30 pl with a

jetting frequency of 33 kHz, while in the QE-256/80 AAA model, it is 80 pl with a jetting frequency of 20 kHz. [13]

2.3.7 FujiFilm Galaxy Series

Galaxy series is a high-performance printhead designed for various industrial and commercial printing and hot melt printing applications at up to 450 dpi resolutions. The series consists of two models that use hot melt colors (Galaxy PH 256/80 HM and Galaxy PH 256/30 HM). In comparison, two models are designed to use classic Inkjet inks (Galaxy PH 256/30 AAA, Galaxy JA 256/80 AAA). The primary characteristic of these heads is the inclusion of 256 nozzles whose diameters are: 52 µm (Galaxy PH 256/80 HM), 42 µm (Galaxy PH 256/30), 36 µm (Galaxy PH 256/30 AAA), 52 µm (Galaxy JA 256 /80 AAA). In this way, with a jetting frequency of 12 kHz and 20 kHz, the size of the droplets was 80 pl and 28 pl. Unlike other versions, the Galaxy JA 256/80 AAA is a compact Inkjet jetting assembly designed specifically for applications that require multiple printheads. [14]

2.3.8 FujiFilm Nova Series

In the category of heads that contain one print line with 256 nozzles, FujiFilm heads of the Nova series are included. The FujiFilm Nova series consists of models PH 256/80 AAA and Nova JA 256/80 AAA intended for a broad range of industrial and commercial printing applications such as wide format graphics, addressing, and packaging. Although they form relatively large droplets (native size 75 pl with a frequency jetting of 20 kHz), connecting them makes it possible to achieve a print resolution of 450 dpi. With an operating temperature range of up to 90°C, applying liquids with a viscosity of 8 mPa·s to 20 mPa·s (aqueous ink, organic solvent ink, UV curable inks) is possible. [15]

2.3.9 FujiFilm S Series

Inkjet FujiFilm series S achieves the lowest reproduction quality. The reason for that is one line of 128 nozzles. This means that higher quality is achieved by combining several heads into one unit. The series consists of 3 models (SE-128 AA, SL-128 AA, SE-128 AA), which differ in the jetting speed and the formed droplets' size. The models of the SE mark thus achieve drops of size 30 pl (jetting frequency 40 kHz). The other two models form larger droplets and layers: SL droplets of 80 pl (jetting frequency with 30 kHz) and SM droplets of 50 pl with a jetting frequency of 20 kHz.
This head type is used when working with liquids containing organic solvents and UV photoinitiators. [16]

2.4 Konica Minolta Inkjet heads

Konica Minolta Inkjet heads are engineered for precision, capable of generating droplets in three driving modes: Shear 3 cycles (Inkjet heads with 512 and 1024 nozzles), Shear independent (Inkjet heads with 512, 1024, and >1024 nozzles), and the band (Inkjet heads 800 and 1600 nozzles).

This precision allows for the formation of three categories of droplets: small (3,5 pl, 6 pl), medium (10 pl, 13 pl, 14 pl), and large (35 pl, 42 pl), with a grey scale (levels) range from 1 to 8.

The principle with the Shear 3-cycle head (models KM512MN, KM1024MNB, KM512LN, KM1024LNB, KM512SHX, M10245HB, KM1024MHB, KM1024LHB) applies exclusively to the input digital waveform signal. The installation of a heater is crucial as it enables the use of UV inks, while the construction without a heater (the first four heads) is only suitable for solvent inks.

The operating principle of Shear independent (KM512iLNB-30, KM1024iSHE, KM1024aSHE, KM1800iSHC-C, KM1024iMHE-D, KM1280MHH, KM1024iLHE-301, KM1024aLHG-RC, KM1024iSAE-C, KM512iMAB-C, KM1024iMAE-C) enables the installation of heaters in Inkjet heads. It enables us to work with all possible inks, as it can receive digital and analog input signals. It is possible to form different volumes of droplets that move in the volume range of 3,5 pl, 6 pl, 10 pl, 13 pl, 14 pl. Based on this, the heads have an output print resolution of 180 dpi, 360 dpi, 450 dpi, and 600 dpi. The "jetting frequency" refers to the rate at which droplets are ejected from the nozzles, and it is 27 kHz, 45 kHz, and 56 kHz for these heads.

Bend Inkjet heads (models KM800, KM1600) use an analog input signal to form multidrop droplets of size from 5 to 22 pl (jetting frequency 18,2 kHz). Heads of this type contain 800 and 1600 nozzles respectively, creating a print resolution of 360 dpi and 720 dpi. They are successfully applied with water-based inks and solvent inks.[17] All these KM heads (as well as the Inkjet head KM1024iLHE-30, which was used in this work) applay piezo crystal's shear movement principle.[18]

3 EXPERIMENTAL PART

In the experiment, we used only the highest quality materials, such as semigloss double-coated cardboard (grammage 300 g/m^2), cated on final format dimensions SRA3+. For print quality evaluation, we used four vertical and four horizontal lines with defined widths (W = 1,40 mm, $W2$ = 1 mm, $W3$ = 0,6 mm, and $W4$ = 0,3 mm. These lines are PDF Fogra line standard test charts used in the printing industry. In the first step, a high-quality reference print was created. The bitmapped vertical and horizontal lines were saved in PDF format (using Adobe Illustrator 2024), making it suitable for the RIP (Hikari KM IC602). This black and white reference imprint characterizes a high-quality print resolution of 1200 × 3600 dpi. For that occasion was used fine black Simitri powder toner and Japanese EP machine Konica Minolta Accurio Print C3080. [20]

The experimental samples were produced on an identical printing substrate. The key machine in the experiment the authors used was a commercial Jetvarnish 3DS film varnishing and printing machine. The main print engine features five Inkjet print heads type KM 1024 LHE-30 (positioned in the configuration of 3 + 2 heads). It also contains a modular UV light source with two LED lamps, which emit a wavelength of 395 nm. The first LED UV lamp has four modules (4×5 W power), while the second LED UV

lamp has one module of 12 W power. Using MGI UV Inkjet curing varnish, three different coating layers were printed (21 µm, 36 µm, and 51 µm). The JETVARNISH 3DS machine featured MGI JVTi 3.3.0.5 software, which adjusts the applied varnish layers. Data on the application of LED UV varnish were not obtained directly (through physical measurement) but indirectly (from the manufacturer's data on the system). The glossy varnish JV3DS contains photoinitiators, acrylic monomers and polymers, and various additives. The main ingredient is the polymer diphenyl (2,4,6-trimethyl benzoyl) and phosphine oxide (CH3)3C6H2COP(O)(C6H5)2). [21]

Murata Gold-type foil was used to create the final gold lines. For this application, alongside the varnish application unit and UV curing, there is a metalized foil application unit. The length of the metalized gold foil was 400 m, meaning a maximum transfer of dimension 32 × 45 cm foil could be achieved. The temperature was controlled in 6 independent zones, with the possibility of changing the foil temperature from 90 °C to 190 °C. In our case, a constant heater temperature of 105 °C was used. MGI JVTi 3.3.0.5 was used to control the application of transparent JV3DS varnish. In contrast, an AIS scanner with 180 defined check positions was used to achieve the exact position of print elements. The varnish channel was created using the alpha channel generated by Adobe Photoshop CS6 25.7. [22]

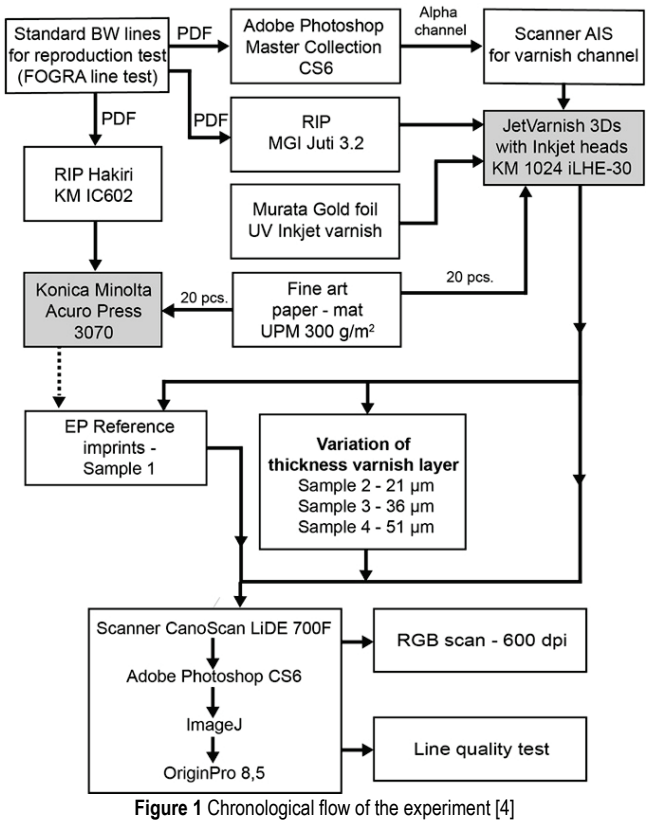


Figure 1 Chronological flow of the experiment [4]

Three different samples were printed, and the thickness of the varnish varied before contact with the metalized foil (21 µm, 36 µm, and 51 µm). After that for line image analysis, the all samples were scanned in a scanner Japan manifacture CanoScan LiDE 700F (RGB image in resoluton of 600 dpi and TIFF format). Before image analysis all line photos were meticulously prepared in Adobe Photoshop to the same TIFF grayscale format (size 9×6 cm with an original resolution of 600 dpi). To ensure the thoroughness and accuracy of our analysis, the final results were obtained using the open-source program ImageJ version 1.5.3 with a carefully designed macro program to calculate line surface coverage.

```
run("HSB Stack");
run("Stack to Images");
selectWindow("Saturation"); close();
selectWindow("Hue");
close();
makeRectangle(9, 6, 456, 574);
makeRectangle(7, 0, 458, 580);
run("Plot Profile");
run("Select All");
```

Based on the detected black pixels, the grayscale profile and width of the line test were determined. The average results were processed and analyzed using OriginPro 8.5. Fow chart of the experiment is shows a Fig. 1 [4].

4 RESULTS AND DISCUSSION

Printed products used for information transfer may contain elements in horizontal and vertical lines (achromatic illustrations, tables, text, letters, etc.). In this research, we aim to enhance the quality of printed products by analyzing four golden lines created by the previous application of UV LED varnish. Figs. 2 and 3 show profiles of these experimental lines with widths of 1,40 mm, 1 mm, 0,6 mm, and 0,3 mm printed on a Konica Minolta AccurioPrint C3070L electrophotographic machine and JETVARNISH 3DS machine.

When compared to the meticulously precise reference BW print, generated by the ROS laser head, the gold print with a 21 µm varnish layer exhibits a reduction in line width. The mutual difference in line width is as follows: line 1 ΔW_{S1_S2} = 53,18 µm; line 2 ΔW_{S1_S2} = 84,87 µm; line 3 ΔW_{S1_S2} = 67,69 µm; line 4 ΔW_{S1_S2} = 88,89 µm. The line with the smallest width experiences the most significant deviation (reduced by 26%), while the widest line has the least deviation (reduced by 3,8%). The total width of all reproduced horizontal lines, created by applying a 21 µm varnish layer and accepting metallic foil, is reduced by ΔW_{sum} = 294,63 µm compared to the reference.

By analyzing the black horizontal lines of the reference Konica Minolta imprints, a characteristic trend is observed. The peripheral parts exhibit greater grayness (Gr_{L1} = 91,53%, Gr_{L2} = 89,57%, Gr_{L3} = 90,40%, Gr_{L4} = 88%) compared to the central part of the lines (Gr_{L1} = 85%, Gr_{L2} = 84,98%, Gr_{L3} = 84,34%, Gr_{L4} = 82,72%). Reduced line width correlates with a more pronounced decrease in edge grayness, offering key insights into their relationship.

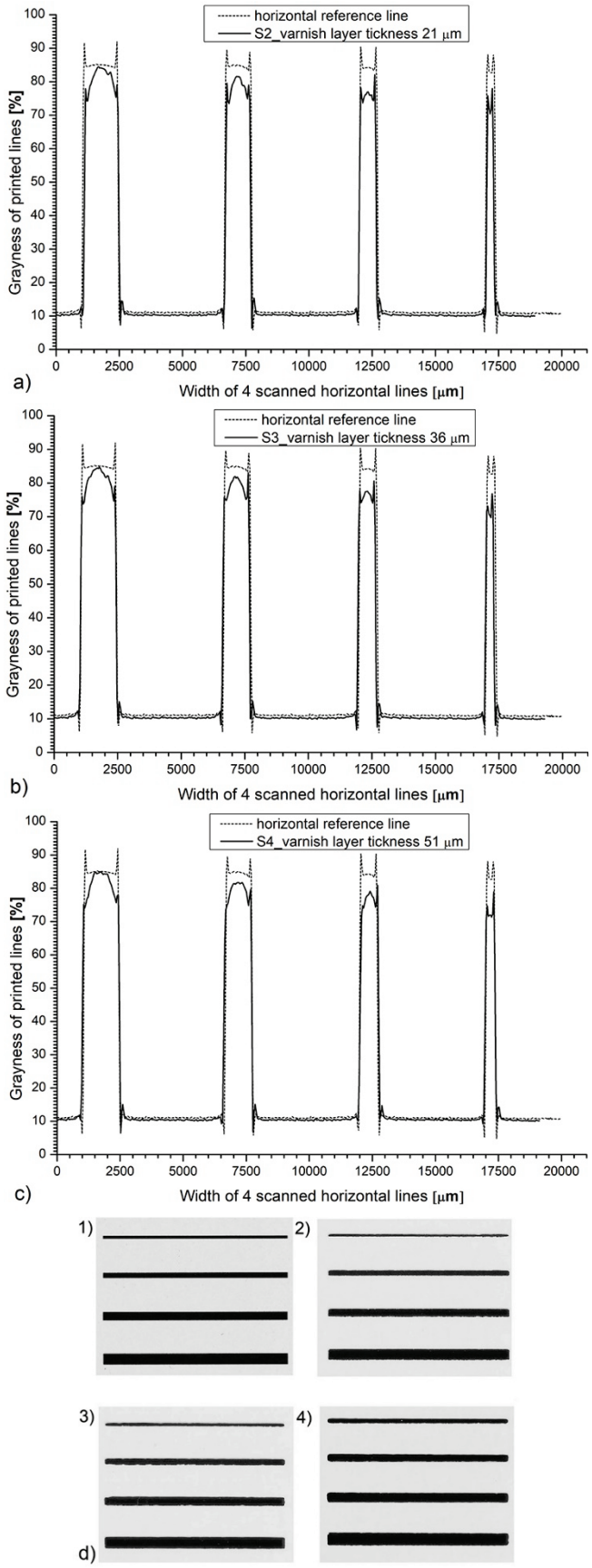


Figure 2 Reproductions of EP reference and four gold horizontal lines: a) varnish layer 21 μm; b) varnish layer 36 μm; c) varnish layer 51 μm; d) enlarged samples used for image analysis (1 = ref; 2 = 21 μm; 3 = 36 μm; 4 = 51 μm).

The line profiles of sample 2 are uneven. With the wider lines (line 1 and line 2), grayness is most significant in the central part of the printed line, while the edges are of lower intensity. With lines of smaller widths (line 3 and line 4), grayness is more pronounced on the edge parts of the print, and the middle parts are less gray. By comparing the grayness of the gold lines with the reference, differences of samples were realized: ΔGr_{L1} = 0,75%, ΔGr_{L2} = 3,45%, ΔGr_{L3} = 8,39%, ΔGr_{L4} = 12,30%. Such amounts lead to the conclusion that the metal foil applied to the varnish reduces the possibility of reproducing gray tones. By lowering the wideness of the lines, the chance of reproducing gray decreases, which means there is also a decrease in information and reflection from the gold foil.

On all tested horizontal lines of the gold imprint (created with a layer of 36 μm), increased dimensions are realized compared to the electrophotographic reference. Their mutual difference amounts to line 1 ΔW_{S1_S3}=111,3 μm; line 2 ΔW_{S1_S3}=73,64 μm; line 3 ΔW_{S1_S3}=87,14 μm; line 4 ΔW_{S1_S3}= 47,73 μm. Thus, of all the printed lines, the widest line will have the smallest deviations (increase of 8%), while the line of smaller width will have the most significant deviation (increase of 13%). With the reference, the total wideness of all reproduced horizontal lines (created by applying varnish with a thickness of 36 μm after subsequent bonding of the metallic foil) increases by ΔW_{sum}=319,81 μm. By reducing the wideness of the lines, the possibility of reproducing gray decreases slightly, but the reproduction is marginally better than with sample 2.

Profiles of horizontal lines on sample 3 are uneven. This is visible in wider lines (line 1, line 2, and line 3), where the grayness is most significant in the center of the printed lines, while the edges are porous and of lower intensity. In the case of the line of smaller width (line 4), the grayness is more pronounced on the edge parts of the print, while the middle parts of the print are of a lesser intensity of grayness. By comparing the average values of the lines of the reference imprint and the average values of the lines printed with foil (sample 3), the following deviations in gray were obtained: ΔGr_{L1} = 0,91%, ΔGr_{L2} = 3,43%, ΔGr_{L3} = 5,37%, ΔGr_{L4} = 10,31%.

Compared to the accurate reference, a gold imprint created with a set-up layer of 51 μm will drastically increase all four horizontal lines. Their mutual difference now amounts to line 1 ΔW_{S1_S4} = 225,63 μm; line 2 ΔW_{S1_S4} = 186,39 μm; line 3 ΔW_{S1_S4} = 226,67 μm; line 4 ΔW_{S1_S4} = 225,52 μm. Thus, regardless of the native width, the three printed horizontal lines will have an identical increment, which means their proportional increase ($L1$ = 16%, $L3$ = 32%, $L4$ = 66%).

A characteristic deviation will be a line with a width of 1 mm, the increase of which is 18%. The total wideness of all reproduced horizontal lines (created by applying varnish with a thickness of 51 μm with the subsequent acceptance of metallic foil) increases by ΔW_{sum} = 864,21 μm compared to the reference.

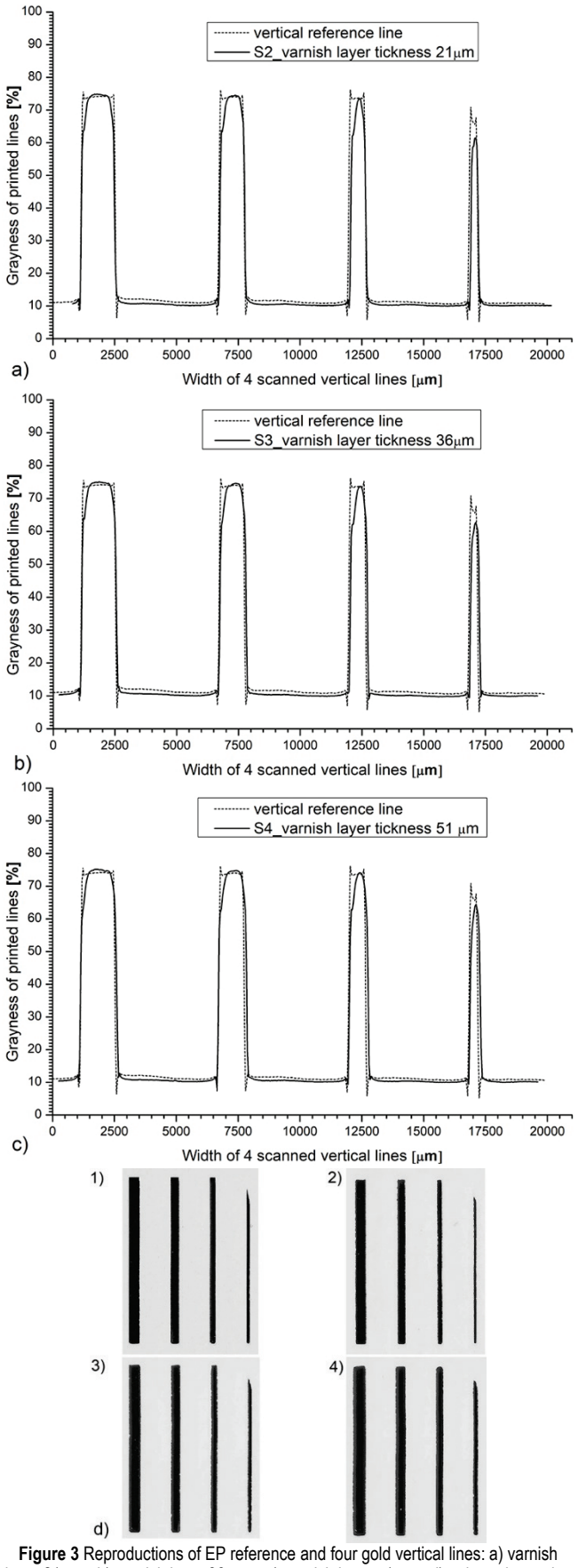


Figure 3 Reproductions of EP reference and four gold vertical lines: a) varnish layer 21 μm; b) varnish layer 36 μm; c) varnish layer 51 μm; d) enlarged samples used for image analysis (1 = ref; 2 = 21 μm; 3 = 36 μm; 4 = 51 μm).

The horizontal line profiles of sample 4 are very uniform. The centers of the line element have a more significant oscillation of the gray value compared to the edge parts of the line. With a line of smaller width (line 4), the grayness of the center and edge parts is almost the same. By comparing the central values of the lines of the reference print and the central values of the lines of the printed lines of sample 4, the difference in grayness of the samples is reduced compared to the previous experimental samples. With line 1, the difference is negligible, while the others are still obvious: ΔGr_{L1} = 0,04%, ΔGr_{L2} = 1,70%, ΔGr_{L3} = 2,96%, ΔGr_{L4} = 7,19%. The results of this research confirm that with a more extensive application of varnish, a better possibility of reproducing the line and its more excellent contrast is realized.

When analyzing the reference vertical line and the line created by applying varnish with a thickness of 21 μm, a trend of decreasing wideness of the golden printed lines can be observed. Their mutual difference amounts to line 1 ΔW_{S1_S2} = 66,36 μm; line 2 ΔW_{S1_S2} = 30,09 μm; line 3 ΔW_{S1_S2} = 90,15 μm; line 4 ΔW_{S1_S2} = 29,73 μm.

This demonstrates the precision of the vertical line printing, with the three widest lines achieving a negative movement (reduction of line width 0,6 mm by 14%), while the line of smaller width will have positive deviations (increase of line wideness 0,3 mm by 14%). The reduction in the total wideness of all 4 sample lines amounts to ΔW_{sum} = 156,87 μm. In relation to all horizontal lines (ΔW_{avgr} = 73,65 μm), the vertical direction will give a smaller deviation in line reproduction (ΔW_{avg} = 54,08 μm). Therefore, the vertical direction with this drift will be slightly more precise.

By analyzing the vertical black lines of the reference print, it is also possible to observe the characteristic phenomenon that the edge parts are of greater grayness (Gr_{L1} = 75,70%, Gr_{L2} = 76,17%, Gr_{L3} = 76,24%, Gr_{L4} = 70,93%) concerning the central part of the lines (Gr_{L1} = 74,20%, Gr_{L2} = 74,07%, Gr_{L3} = 73,70%, Gr_{L4} = 65,61%). However, there is no significant grayness reduction here (except for line 4). The vertical lines' maximum grayness is smaller than the horizontal lines. This shows that the direction of printing (orientation of lines) affects the quality of line reproduction; that is, the laser head and the application of tones have different effects on the quality of line realization.

Profiles of vertical lines on sample 2 are regular, and there are no significant deviations, which is the case with profiles of horizontal lines. By comparing the average gray values of the lines of the reference print and the average gray values of sample 2, it is noticed that the gold foil is accepted on the varnish in almost the exact amounts. The exception is the thinnest line, where the grayness is reduced by ΔGr_{L4} = 4,18%. A more significant reduction of depo-sition in the edge zone of the lines was also observed. Thus, the beginning of the line will be slightly less reproduced. By comparing horizontal and vertical lines, it was noticed that the vertical distribution of lines would give a better and higher quality golden line. This is partly caused by a configuration where the modular head is stationary and the substrate moves during printing. This precision is evident in the minimal deviation in the line reproduction.

When analyzing the reference vertical line and the line created by applying varnish with a thickness of 36 µm, a trend of increasing the wideness of the golden printed lines can be observed. Their mutual difference amounts to line 1 ΔW_{S1_S3} = 116,57 µm; line 2 ΔW_{S1_S3} = 119,25 µm; line 3 ΔW_{S1_S3} = 59,37 µm; line 4 ΔW_{S1_S3} = 185,68 µm. Thus, of all the printed lines, the largest deviation will be the line of 0,3 mm (an increase of 86 %), and the smallest deviation of the lines will be 0,6 mm (an increase of 9%). The increase in the total wideness of all 4 sample lines amounts to ΔW_{sum} = 480,87, µm. With all horizontal lines (ΔW_{avg} = 79,95 µm), the vertical direction will give a more significant deviation in the reproduction of lines (ΔW_{avg} = 120,21 µm) with an increase in varnish lađer of 15 µm. Therefore, the vertical direction with this application will not be more precise in terms of dimensions.

The line profiles of sample 3 are regular, and there are no major deviations as in the horizontal line profiles. This means that the maximum amounts of grayness are in the area of the center of the line. By comparing the average value of the lines of the reference print and the average values of the lines printed with the foil of sample 3, a more pronounced deviation occurs only in line 4. The same anomaly is also observed in sample 3, where the amounts of grayness are slightly higher than the reference sample (ΔGr_{L1} = 0,61%, ΔGr_{L2} = 0,79%, ΔGr_{L3} = 0,35%, ΔGr_{L4} = 1,51%). This leads to the conclusion that the grayscale reproduction of the vertical lines of sample 4 is also almost identical to the reference sample.

The vertical line of sample 4 (with varnish coating thick 51 µm with the subsequent application of metallic foil) will achieve a significant increase in line thickness compared to the reference lines. Their mutual difference amounts to: line 1 ΔW_{S1_S4} = 337,2 µm; line 2 ΔW_{S1_S4} = 224,88 µm; line 3 ΔW_{S1_S4} = 191 µm; line 4 ΔW_{S1_S4} = 307,21 µm. This further demonstrates the precision of the vertical line printing, with the line of 0,3 mm (increase by 142%) having the largest deviation, and the line having the smallest deviation of 1 mm (increase by 22%). Compared to the reference sample, the total wideness (all 4 lines) increases by ΔW_{sum} = 1060,42 µm. The mean value of the change in the wideness of the lines with an increase in varnish layer of 30 µm is ΔW_{avg} = 265,11 µm. This alone continues the trend that the vertical direction is more imprecise (gives slightly larger line impressions).

The profiles of the vertical lines of sample 4 are regular, and there are no deviations like the profiles of the horizontal lines. This means the maximum amounts of gray are in the cent area of the vertical line printing element. By comparing the average values of the lines of the reference print and the average values of the lines of printed sample 4, there are no significant deviations in contrast (ΔGr_{L1} = 0,79%, ΔGr_{L2} = 0,89%, ΔGr_{L3} = 0,69%, ΔGr_{L4} = 0,78%). Namely, the possibility of reproducing gray is increasing, and the possibility of reproducing sample 4 is better than that of the reference sample.

This results shows that the varnish application is not optimal and need aditional corection in printing file. An optimal metallized print on semigloss cardboard could be achieved by a targeted reduction of line thickness depending on the height of the varnish application (the higher the varnish application, the lower the speed of the machine, which affects the spreading of the UV varnish).

5 CONCLUSION

The optimal experimental setup for the KM Inkjet head KM1024iLHE-30 involves applying a 21 µm UV drying varnish on paper. This setup is best for achieving accurate lines for protective printing and printing with strict metal conductors.

Increasing the varnish layer by 15 µm or 30 µm will significantly alter the dimensions of the realized lines, possibly distorting the image and text.

The quality of the realized lines is affected by the direction (verticals and horinzontals) of line in orginal files. With minimal varnish application, vertical lines are more precise than horizontal ones. However, as the varnish layer increases, the width of the vertical lines increases more significantly than the horizontal lines. Increasing the varnish amount leads to growing increments, resulting in less precise line imprints in the cross direction than in the machine direction.

The line profile reflects the homogeneity of line reproduction. Increasing the varnish application achieves better uniformity of lines, with the central part being more contrasting than the peripheral part. Additionally, wider lines will exhibit more pronounced gray values.

Changing the printing direction will impact the realized grayness. Vertical lines (print-direction) produce more precise and homogeneous lines. Lines with a width of 0,3 mm will have the most significant oscillations in width, while a 1 mm line will have the least oscillations during printing.

To achieve optimal results, an Inkjet head with a smaller print nozzle and volume will be necessary. Further research will focus on testing of volume of applied UV varnish and adopting referent file to achieve better outcomes in realizing targeted printing elements. Therefore, it is planned to test the possibility of using a larger number of print heads and higher printing speeds.

6 REFERENCES

[1] Majnarić, I. (2016). *Osnove digitalnog tiska*. Zagreb, Hrvatska: Sveučilište u Zagrebu Grafički fakultet. (in Croatian)

[2] Kim, Y. Sim, W. Park C., Yoo Y., Joung J., & Oh Y. (2006). The Effects of Driving Waveform of Piezoelectric Industrial Inkjet Head for Fine Patterns. *1st IEEE International Conference on Nano/Micro Engineered and Molecular Systems,* 826–831 doi: 10.1109/NEMS.2006.334905

[3] Chaffe-Stengel, P., & Stengel, D. (2013). Future, Opportunities and Challenges of Inkjet; *Technologies. Atomization and Sprayst, 23*(6). 1–13. doi: 10.1615/AtomizSpr.2013007653.

[4] Majnarić I., Morić M., Valdec D., & Milković K., (2024). The Effect of Applying UV LED-Cured Varnish to Metalized Printing Elements during Cold Foil Lamination, *Coatings 14* (604). 1–12. doi: 10.3390/coatings14050604.

[5] Schozer. R. (2024, May 27). Technology Outlook: Digital printing. What They Think? Retrieved from:

https://d3a577syzx0or3.cloudfront.net/webinars/2023-technology-outlook-digital-print.pdf.
[6] Kyocera Corporation (2024, May 23). Inkjet head Kyocera, Brochure, 2024, Retrieved from: https://global.kyocera.com /prdct/inkjetprintheads/application/detail/index.html.
[7] Ricoh Corporation (2024, June 15). Inkjet head Ricoh, Brochure, 2024, Retrieved from: https://industry.ricoh.com /Media/Ricoh/Sites/industry/industrialinkjet/pdf/RICOH_GH 2220.pdf.
[8] FUJIFILM Corporation (2024, July 8). Technical data FUJIFILM Dimatix SKYFIRE printhead, Brochure, 2024, Retrieved from: https://www.fujifilm.com/us/en/business/ inkjetsolutions/industrial-printheads/skyfireSF600/features.
[9] FUJIFILM Corporation (2024, July 8). Technical data FUJIFILM Dimatix SAMBA printhead, Brochure, Brochure, 2024, Retrieved from: https://www.fujifilm.com /us/en/business/inkjetsolutions/industrial-printheads /sambasf3GL/features.
[10] FUJIFILM Corporation (2024, July 9). Technical data FUJIFILM Dimatix STARFIRE printhead, Brochure, 2022, Retrieved from: https://www.fujifilm.com/us/en/business/ inkjet-solutions/industrial-printheads/StarFireSG600-A-Series /features.
[11] FUJIFILM Corporation (2024, July 8). Technical data FUJIFILM Dimatix POLARIS printhead, Brochure, 2020, Retrieved from: https://www.fujifilm.com/us/en/business/ inkjet-solutions/industrialprintheads/polarisPDS00071 /features.
[12] FUJIFILM Corporation (2024, July 11). Technical data FUJIFILM Dimatix SAPPHIRE printhead, Brochure, 2022, Retrieved from: https://www.fujifilm.com/us/en/business/ inkjetsolutions/industrialprintheads/SapphireQSPrintheads/fea tures.
[13] FUJIFILM Corporation (2024, July 11). Technical data FUJIFILM Dimatix EMERALD printhead, Brochure, 2022, Retrieved from: https://www.fujifilm.com/us/en/business/ inkjetsolutions /industrial-printheads/PDS00120/features.
[14] FUJIFILM Corporation (2024, July 8). Technical data FUJIFILM Dimatix GALAXY printhead, Brochure, 2020, Retrieved from: https://www.fujifilm.com/us/en/business/ inkjetsolutions/industrial-printheads/PDS00014/features
[15] FUJIFILM Corporation (2024, July 11). Technical data FUJIFILM Dimatix NOVA printhead, Brochure, 2020, Retrieved from: https://www.fujifilm.com/us/en/business/ inkjetsolutions/industrial-printheads/NovaPH256/80AAA/ features
[16] FUJIFILM Corporation (2024, July 11). Technical data FUJIFILM Dimatix S SERIE printhead, Brochure, 2020. Retrieved from: https://www.fujifilm.com/us/en/business/ inkjetsolutions/industrial-printheads /SM-128 AA/features
[17] Konica Minolta (2024, June 22). Inkjet Printhead Lineup Catalogs, Retrieved from: https://www.konicaminolta.com /globalen/inkjethead/products/inkjethead/pdf/ij_lineup_jpn_en g.pdf
[18] Konica Minolta (2024, June 22). Material Safety Data Sheet Industrial Inkjet KM1024 Series, Retrieved from: https://www.konicaminolta.com/globalen/inkjethead/products /inkjethead/pdf/ij_lineup_jpn_eng.pdf
[19] Konica Minolta (2024, June 22). AccurioPress C3080, Retrieved from: https://www.konicaminolta.eu/en/ businesssolutions/products/production-printing/colour /accuriopress-c30703080/introduction.html
[20] UPM (2022, September 21). UPM Finesse silk, Retrieved from: https://www.upmpaper.com/products/papercatalogue/ categories/sheet-fed-offset-papers/upm-finesse-silk/
[21] Konica Minolta (2023, May 12). Material Safety Data Sheet for JV3DS - LED Varnish, Retrieved from: https://www.konicaminolta.jp/about/csr/environment/commun ication/msds/pdf/mgi/MGI_SDS-MSDS_n°58-JV3DS-LED_VARNISH_PN_10143S_020522.pdf
[22] Konica Minolta (2024, June 22). AIS SmartScaner, Retrieved from: https://www.printingnews.com/softwareworkflow/ product/12252965/mgi-digital-graphic-technology-ais-smartscanner

Authors' contacts:

Igor Majnarić, Full. Prof. PhD
University of Zagreb Faculty of Graphic Arts,
Getaldićeva 2, 10000 Zagreb, Croatia
Igor.majnaric@grf.unizg.hr

Marko Morić, Asst. PhD
(Corresponding author)
University North,
Trg Žarka Dolinara 1, 48000 Koprivnica, Croatia
mmoric@unin.hr

Marko Kadija, mag. ing. techn. graph.
Konica Minolta Croatia - Business Solutions, Ltd.,
Horvatova 82, 10000 Zagreb, Croatia
marko.kadija@konicaminolta.hr

Krunoslav Hajdek, Assoc. Prof. PhD
University North,
Trg Žarka Dolinara 1, 48000 Koprivnica, Croatia
khajdek@unin.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20.....

Categorization of paper
Received: y-m-d, Accepted: y1-m1-d1

14pt
14pt

Article Title Only in English (Style: Arial Narrow, Bold, 14pt)

14pt

Ivan Horvat, Thomas Johnson, Marko Marić (Style: Arial Narrow, Normal, 10pt)

14pt

Abstract: Article abstract contains maximum of 150 words and is written in the language of the article. The abstract should reflect the content of the article as precisely as possible. TECHNICAL JOURNAL is a trade journal that publishes scientific and professional papers from the domain(s) of mechanical engineering, electrical engineering, civil engineering, multimedia, logistics, etc., and their boundary areas. This document must be used as the template for writing articles so that all the articles have the same layout. (Style: Arial Narraw, 8pt)

8pt

Keywords: keywords in alphabetical order (5-6 key words). Keywords are generally taken from the article title and/or from the abstract. (Style: Arial Narraw, 8pt)

10pt
10pt

1 INTRODUCTION (Article Design)

(Style: Arial Narrow, Bold, 10pt)

10pt

(Tab 6 mm) The article is written in Latin script and Greek symbols can be used for labelling. The length of the article is limited to eight pages of international paper size of Letter (in accordance with the template with all the tables and figures included). When formatting the text the syllabification option is not to be used.

10pt

1.1 Subtitle 1 (Writing Instructions)

(Style: Arial Narrow, 10pt, Bold, Align Left)

10pt

The document format is Letter with margins in accordance with the template. A two column layout is used with the column spacing of 10 mm. The running text is written in Times New Roman with single line spacing, font size 10 pt, alignment justified.

Article title must clearly reflect the issues covered by the article (it should not contain more than 15 words).

Body of the text is divided into chapters and the chapters are divided into subchapters, if needed. Chapters are numbered with Arabic numerals (followed by a period). Subchapters, as a part of a chapter, are marked with two Arabic numerals i.e. 1.1, 1.2, 1.3, etc. Subchapters can be divided into even smaller units that are marked with three Arabic numerals i.e. 1.1.1, 1.1.2, etc. Further divisions are not to be made.

Titles of chapters are written in capital letters (uppercase) and are aligned in the centre. The titles of subchapters (and smaller units) are written in small letters (lowercase) and are aligned left. If the text in the title of the subchapter is longer than one line, no hanging indents.

10pt

Typographical symbols (bullets), which are being used for marking an item in a list or for enumeration, are placed at a beginning of a line. There is a spacing of 10pt following the last item:

- Item 1
- Item 2
- Item 3

10pt

The same rule is valid when items are numbered in a list:

1) Item 1
2) Item 2
3) Item 3

10pt

1.2 Formatting of Pictures, Tables and Equations

(Style: Arial Narrow, 10pt, Bold, Align Left)

10pt

Figures (drawings, diagrams, photographs) that are part of the content are embedded into the article and aligned in the centre. In order for the figure to always be in the same position in relation to the text, the following settings should be defined when importing it: text wrapping / in line with text.

Pictures must be formatted for graphic reproduction with minimal resolution of 300 dpi. Pictures downloaded from the internet in ratio 1:1 are not suitable for print reproduction because of unsatisfying quality.

10pt

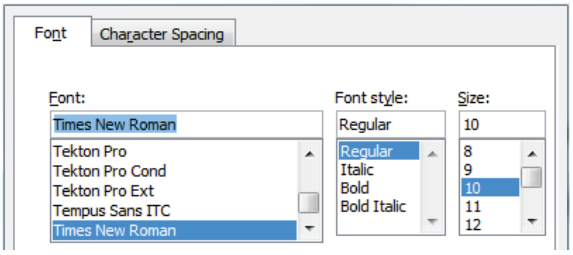


Figure 1 Text under the figure [1]
(Style: Arial Narraw, 8pt, Align Centre)

10pt

The journal is printed in black ink and the figures have to be prepared accordingly so that bright tones are printed in a satisfactory manner and are readable. Figures are to be in colour for the purpose of digital format publishing. Figures in the article are numbered with Arabic numerals (followed by a period).

Text and other data in tables are formatted - Times New Roman, 8pt, Normal, Align Center.

When describing figures and tables, physical units and their factors are written in italics with Latin or Greek letters, while the measuring values and numbers are written upright.

10pt

Table 1 Table title aligned centre
(Style: Arial Narraw, 8pt, Align Centre)

	1	2	3	4	5	6
ABC	ab	ab	ab	ab	ab	ab
DEF	cd	cd	cd	cd	cd	cd
GHI	ef	ef	ef	ef	ef	ef

10 pt

Equations in the text are numbered with Arabic numerals inside the round brackets on the right side of the text. Inside the text they are referred to with equation number inside the round brackets i.e. ".... from Eq. (5) follows ...." (Create equations with MathType Equation Editor - some examples are given below).

10pt

$$F_{\text{avg}}(t, t_0) = \frac{1}{t}\int_{t_0}^{t_0+t} F\left[q(\tau), p(\tau)\right] \mathrm{d}\tau, \tag{1}$$

$$\cos\alpha + \cos\beta = 2\cos\frac{\alpha+\beta}{2}\cdot\cos\frac{\alpha-\beta}{2}, \tag{2}$$

$$(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}, \tag{3}$$

$$AAMC = \frac{1}{n}\sum_{i=1}^{n} PVMC_i. \tag{4}$$

10pt

Variables that are used in equations and also in the text or tables of the article are formatted as *italics* in the same font size as the text.

10pt

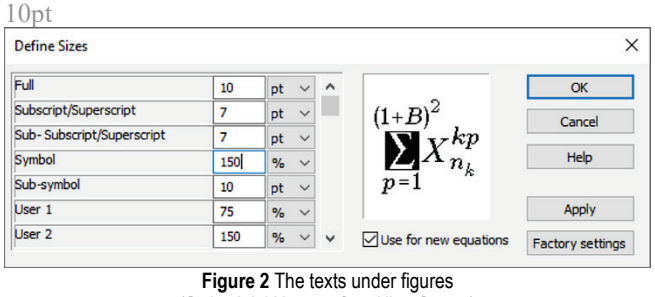


Figure 2 The texts under figures
(Style: Arial Narraw, 8pt, Align Centre)

10pt

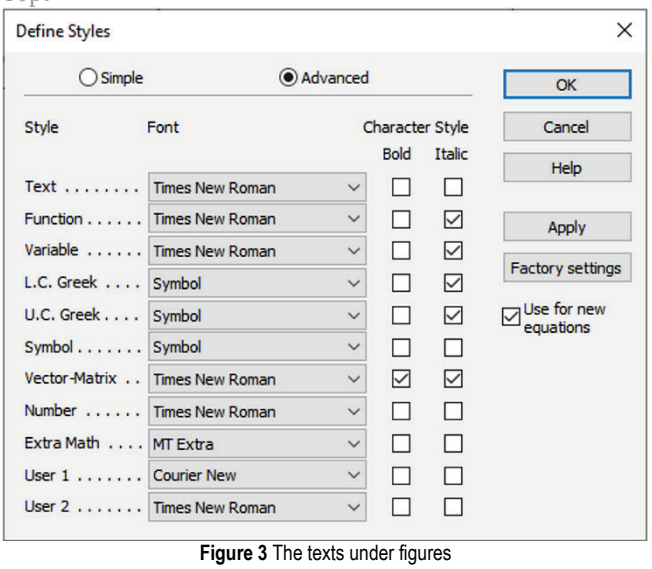


Figure 3 The texts under figures
(Style: Arial Narraw, 8pt, Align Centre)

Figures and tables that are a part of the article have to be mentioned inside the text and thus connected to the content i.e. " … as shown in Fig. 1…" or "data from Tab. 1…" and similar.

10pt

2 PRELIMINARY ANNOTATION

10pt

Article that is offered for publication cannot be published beforehand, be it in the same or similar form, and it cannot be offered at the same time to a different journal. Author or authors are solely responsible for the content of the article and the authenticity of information and statements written in the article.

Articles that are accepted for publishing are classified into four categories: original scientific papers, preliminary communications, subject reviews and professional papers.

Original scientific papers are articles that according to the reviewer and the editorial board contain original theoretical or practical results of research. These articles need to be written in such a way that based on the information given, the experiment can be repeated and the results described can be achieved together with the author's observations, theoretical statements or measurements.

Preliminary communication contains one or more pieces of new scientific information, but without details that allow recollection as in original scientific papers. Preliminary communication can give results of an experimental research, results of a shorter research or research in progress that is deemed useful for publishing.

Subject review contains a complete depiction of conditions and tendencies of a specific domain of theory, technology or application. Articles in this category have an overview character with a critical review and evaluation. Cited literature must be complete enough to allow a good insight and comprehension of the depicted domain.

Professional paper can contain a description of an original solution to a device, assembly or instrument, depiction of important practical solutions, and similar. The article need not be related to the original research, but it should contains a contribution to an application of known scientific results and their adaptation to practical needs, so it presents a contribution to spreading knowledge, etc.

Outside the mentioned categorization, the Editorial board of the journal will publish articles of interesting content in a special column. These articles provide descriptions of practical implementation and solutions from the area of production, experiences from device application, and similar.

10pt

3 WRITING AN ARTICLE

10pt

Article is written in the English language and the terminology and the measurement system should be adjusted to legal regulations, standards and the International System of Units (SI) (Quantities and Units: ISO 80 000 - from Part 1 to Part 14). The article should be written in third person.

Introduction contains the depiction of the problem and an account of important results that come from the articles that are listed in the cited literature.

Main section of the article can be divided into several parts or chapters. Mathematical statements that obstruct the reading of the article should be avoided. Mathematical statements that cannot be avoided can be written as one or more addendums, when needed. It is recommended to use an example when an experiment procedure, the use of the work in a concrete situation or an algorithm of the suggested method must be illustrated. In general, an analysis should be experimentally confirmed.

Conclusion is a part of the article where the results are being given and efficiency of the procedure used is emphasized. Possible procedure and domain constraints where the obtained results can be applied should be emphasized.

AI and AI-assisted tools do not qualify for authorship under TG/TJ's authorship policy. Authors who use AI or AI-assisted tools during the manuscript writing process are asked to disclose their use in a separate section of the manuscript. The publishing agreement process works as usual, with the authors keeping the copyright to their own work.

10pt

4 RECAPITULATION ANNOTATION

10pt

In order for the articles to be formatted in the same manner as in this template, this document is recommended for use when writing the article. Finished articles written in MS Word for Windows and formatted according to this template must be submitted using our The Paper Submission Tool (PST) (https://tehnickiglasnik.unin.hr/authors.php) or eventually sent to the Editorial board of the Technical Journal to the following e-mail address: tehnickiglasnik@unin.hr

The editorial board reserves the right to minor redaction corrections of the article within the framework of prepress procedures. Articles that in any way do not follow these authors' instructions will be returned to the author by the editorial board. Should any questions arise, the editorial board contacts only the first author and accepts only the reflections given by the first author.

10pt

5 REFERENCES (According to APA)

10pt

The literature is cited in the order it is used in the article. No more than 35 references are recommended. Individual references from the listed literature inside the text are addressed with the corresponding number inside square brackets i.e. "… in [7] is shown …". If the literature references are web links, the hyperlink is to be removed as shown with the reference number 8. Also, the hyperlinks from the e-mail addresses of the authors are to be removed. In the literature list, each unit is marked with a number and listed according to the following examples (omit the subtitles over the references – they are here only to show possible types of references):

9pt

[1] See http://www.bibme.org/citation-guide/apa/

[2] See http://sites.umuc.edu/library/libhow/apa_examples.cfm

[3] (Style: Times New Roman, 9pt, according to APA)

[4] Amidzic, O., Riehle, H. J. & Elbert, T. (2006). Toward a psychophysiology of expertise: Focal magnetic gamma bursts as a signature of memory chunks and the aptitude of chess players. *Journal of Psychophysiology, 20*(4), 253-258. https://doi.org/10.1027/0269-8803.20.4.253

[5] Reitzes, D. C. & Mutran, E. J. (2004). The transition to retirement: Stages and factors that influence retirement adjustment. *International Journal of Aging and Human Development, 59*(1), 63-84. Retrieved from http://www.baywood.com/journals/PreviewJournals.asp?Id=0091-4150

[6] Jans, N. (1993). *The last light breaking: Life among Alaska's Inupiat Eskimos*. Anchorage, AK: Alaska Northwest Books.

[7] Miller, J. & Smith, T. (Eds.). (1996). *Cape Cod stories: Tales from Cape Cod, Nantucket, and Martha's Vineyard*. San Francisco, CA: Chronicle Books.

[8] Chaffe-Stengel, P. & Stengel, D. (2012). *Working with sample data: Exploration and inference*. https://doi.org/10.4128/9781606492147

[9] Freitas, N. (2015, January 6). People around the world are voluntarily submitting to China's Great Firewall. Why? Retrieved from http://www.slate.com/blogs/future_tense/2015/01/06/tencent_s_wechat_worldwide_internet_users_are_voluntarily_submitting_to.html (Style: Times New Roman, 9pt, according to APA)

10pt

10pt

Authors' contacts:

8pt

Full Name, title
Institution, company
Address
Tel./Fax, e-mail

8pt

Full Name, title
Institution, company
Address
Tel./Fax, e-mail

Note: Gray text should be removed in the final version of the article because it is for guidance only.

Oberwolfach Meetings 2027

Mathematisches Forschungsinstitut Oberwolfach

Member of Leibniz Association

Workshops

Participants of the Workshops at Oberwolfach are invited personally by the Director of the Institute. Participation is subject to such an invitation. Interested researchers, in particular early career scientists, can contact the administration of the Institute. Since the number of participants is restricted, not all inquiries can be considered.

03.01. – 08.01.2027 ID: 2701

Moduli Theory in Differential Geometry and Analysis
B. Hanke (Augsburg), I. Alonso Lorenzo (Piscataway), C. Mantoulidis (Houston), M. Upmeier (Aberdeen)

10.01. – 15.01.2027 ID: 2702

Geometry of Motives
S. Gille (Edmonton), C. Ravi (Mumbai), P. Sechin (Regensburg), A. Vishik (Nottingham)

17.01. – 22.01.2027 ID: 2703

Wave-Kinetic Interfaces: Dispersive Dynamics Meets Kinetic Theory
M. Iacobelli (Zürich), N. Pavlovic (Austin), G. Staffilani (Cambridge MA), N. Visciglia (Pisa)

24.01. – 29.01.2027 ID: 2704

Geometric Methods in Data Science
M. Bauer (Tallahassee), X. Pennec (Sophia Antipolis), S. Sommer (København), G. Steidl (Berlin)

31.01. – 05.02.2027 ID: 2705

Non-commutative Function Theory: A Link between Operator Theory and Random Matrices
M. Banna (Abu Dhabi), G. Cébron (Toulouse), M. Hartz (Saarbrücken), J. E. McCarthy (St Louis)

07.02. – 12.02.2027 ID: 2706a

Mahler Measures and L-Functions
M. Lalín (Montréal), D. van Straten (Mainz), W. Zudilin (Nijmegen)

07.02. – 12.02.2027 ID: 2706b

Lagrangian Torus Fibrations
M. Abouzaid (Stanford), E. Mazzon (Paris), H. Ruddat (Stavanger)

14.02. – 19.02.2027 ID: 2707

Scientific Machine Learning: Mathematical Foundations and Algorithmic Advances
M. Han Veiga (Columbus), A. Heinlein (Delft), C. B. Schönlieb (Cambridge UK), J. Zech (Heidelberg)

21.02. – 26.02.2027 ID: 2708

Hypoelliptic Operators in Geometry
W. Bauer (Hannover), X. Ma (Tianjin/Paris), N. Savale (Dublin), Y. Wang (Baltimore)

28.02. – 05.03.2027 ID: 2709

Statistical Analysis of Black Box Models
P. Bartlett (Berkeley), S. Langer (Bochum), J. Rousseau (Paris), S. Wang (Berlin)

Mini-Workshops

01.08. – 06.08.2027 (2731a/b)
deadline for appl.: 31.01.2027

08.08. – 13.08.2027 (2732a/b)
deadline for appl.: 31.01.2027

15.08. – 20.08.2027 (2733a/b)
deadline for appl.: 31.01.2027

22.08. – 27.08.2027 (2734a/b)
deadline for appl.: 31.01.2027

Time slots for four further Mini-Workshops in winter 2027 will be announced on our website in 2026.

Tandem- or Mini-Workshops

05.09. – 10.09.2027 (2736a/b)
deadline for appl.: 01.09.26/31.01.27

Oberwolfach Seminars

16.05. – 21.05.2027 (2720a/b)
to be announced in winter 2026

24.10. – 29.10.2027 (2743a/b)
to be announced in winter 2026

21.11. – 26.11.2027 (2747a/b)
to be announced in winter 2026

Arbeitsgemeinschaft (study group)

14.03. – 19.03.2027 (2711)
to be announced in spring 2026

10.10. – 15.10.2027 (2741)
to be announced in fall 2026

17.10. – 22.10.2027 (2742),
to be announced in fall 2026

Proposals invited

07.03. – 12.03.2027 ID: 2710

Representations of Finite Groups
O. Dudas (Marseilles), R. Kessar (Manchester), B. Späth (Wuppertal), P. H. Tiep (Rutgers)

21.03. – 26.03.2027 ID: 2712

Homotopy Theory
J. Grodal (København), M. Hill (Minneapolis), B. Richter (Hamburg)

28.03. – 02.04.2027 ID: 2713

Generalized Computability Theory
J. P. Aguilera (Wien), N. Greenberg (Wellington), B. Löwe (Hamburg/Cambridge UK), M. I. Soskova (Madison)

04.04. – 09.04.2027 ID: 2714

Aperiodic Order: Spectral Theory, Harmonic Analysis and Dynamical Systems in Interaction
D. Damanik (Houston), D. Lenz (Jena), N. P. Frank (Poughkeepsie), R. Yassawi (London)

Check online!

11.04. – 16.04.2027 ID: 2715

Geometric, Algebraic and Topological Combinatorics
M. Juhnke (Osnabrück), E. Nevo (Jerusalem/Valladolid), F. Santos (Santander), V. Welker (Marburg)

18.04. – 23.04.2027 ID: 2716

Structure Preserving Methods for Computational Geometric Mechanics
E. Celledoni (Trondheim), S. Ober-Blöbaum (Paderborn), A. Palha (Delft), A. Stern (St Louis)

25.04. – 30.04.2027 ID: 2717

Mathematical Analysis of Fluid Dynamics Models
I. Kukavica (Los Angeles), A. Mazzucato (University Park), E. S. Titi (Cambridge UK), E. Wiedemann (Erlangen)

02.05. – 07.05.2027 ID: 2718

Statistical Physics and Random Surfaces
C. Garban (Lyon), J. Miller (Cambridge UK), S. Sheffield (Cambridge MA)

09.05. – 14.05.2027 ID: 2719

Complexity Theory
M. Bläser (Saarbrücken), I. Dinur (Princeton), R. O'Donnell (Pittsburgh), S. Vadhan (Harvard)

23.05. – 28.05.2027 ID: 2721

Interactions between Algebraic Geometry and Non-commutative Algebra
W. Crawley-Boevey (Bielefeld), O. Schiffmann (Orsay), S. Spenko (Bruxelles), M. Wemyss (Glasgow)

30.05. – 04.06.2027 ID: 2722a

Nonlinear Optics: Physics, Analysis, and Numerics
K. Busch (Berlin), R. El-Ganainy (St Louis), M. Hochbruck (Karlsruhe), C. Lasser (München)

30.05. – 04.06.2027 ID: 2722b

Methods of Mathematization around 1900
H. Junker (Bonn), D. Kent (St Andrews), J. Lorenat (Claremont)

06.06. – 11.06.2027 ID: 2723

Arithmetic of Shimura Varieties
K. Madapusi (Chestnut Hill), T. Richarz (Darmstadt), E. Viehmann (Münster), M. Zhang (Princeton)

13.06. – 18.06.2027 ID: 2724

Homological Commutative Algebra and Applications to Geometry
H. Brenner (Osnabrück), D. Eisenbud (Berkeley), E. Faber (Graz), C. Miller (Syracuse)

20.06. – 25.06.2027 ID: 2725

Recent Trends in Algebraic Geometry
G. Farkas (Berlin), M. Mustata (Ann Arbor), R. Pandharipande (Zürich), C. Voisin (Paris)

27.06. – 02.07.2027 ID: 2726

Mathematical Horizons in Artificial Intelligence
C. Boyer (Paris), M. Hein (Tübingen), R. Nowak (Madison), F. Yang (Zürich)

04.07. – 09.07.2027 ID: 2727

Differentialgeometrie im Grossen
O. Chodosh (Stanford), J. Lotay (Oxford), A. Lytchak (Karlsruhe), L. Wang (New Haven)

11.07. – 16.07.2027 ID: 2728

Dynamische Systeme
S. Crovisier (Gif-sur-Yvette), U. L. Hryniewicz (Aachen), T. M-Seara (Barcelona)

18.07. – 23.07.2027 ID: 2729

Discrete Nonlinear Optimization: Mathematics and Algorithms
J. Lee (Ann Arbor), R. Sotirov (Tilburg), A. Wiegele (Klagenfurt)

25.07. – 30.07.2027 ID: 2730

Partial Differential Equations
A. Fraser (Vancouver), T. Lamm (Karlsruhe), C. Mantoulidis (Houston), X. Ros-Oton (Barcelona)

29.08. – 03.09.2027 ID: 2735

Mathematical Aspects of General Relativity
C. Cederbaum (Tübingen), G. Holzegel (Münster), J. Luk (Stanford), T. Oliynyk (Clayton)

12.09. – 17.09.2027 ID: 2737

Homotopical Algebra and Higher Structures
T. Dyckerhoff (Hamburg), J. Holstein (Hamburg), M. Livernet (Paris), M. Markl (Prague)

19.09. – 24.09.2027 ID: 2738

Proof Complexity and Beyond
P. Beame (Seattle), M. Mahajan (Chennai/Mumbai), J. Nordström (Copenhagen/Lund), P. Pudlák (Praha)

26.09. – 01.10.2027 ID: 2739

Quantum Groups – Algebra, Analysis and Category Theory
S. Neshveyev (Oslo), D. Nikshych (Durham NH), J. Plavnik (Bloomington), A. Skalski (Warsaw)

03.10. – 08.10.2027 ID: 2740

Polytopal Finite Element Methods for PDEs: Theory, Algorithms, and Machine Learning Enhancement
P. Antonietti (Milano), L. Beirão da Veiga (Milano), J. Droniou (Montpellier), P. Wriggers (Hannover)

31.10. – 05.11.2027 ID: 2744

New Directions in Tomography: Theory and Applications
G. Ambartsoumian (Arlington), S. Arridge (London), B. Hahn-Rigaud (Stuttgart), A. Wald (Göttingen)

07.11. – 12.11.2027 ID: 2745

Applications of Optimal Transportation
G. Carlier (Paris), E. Radici (L'Aquila), F. Santambrogio (Lyon), B. Schmitzer (Göttingen)

14.11. – 19.11.2027 ID: 2746

Combinatorial Optimization
J. de Loera (Davis), V. Traub (Zürich), L. Végh (Bonn)

28.11. – 03.12.2027 ID: 2748

Applied Harmonic Analysis and Data Science
R. Alaifari (Aachen), A. Bandeira (Zürich), F. Krahmer (Garching), R. Ward (Austin)

05.12. – 10.12.2027 ID: 2749

Machine Learning in Infinite Dimensions
M. de Hoop (Houston), Y. Korolev (Bath), N. Mücke (Braunschweig), M. Thorpe (Coventry)

12.12. – 17.12.2027 ID: 2750

Convex Geometry and its Applications
S. Artstein-Avidan (Tel Aviv), D. Hug (Karlsruhe), E. Werner (Cleveland)



www.mfo.de/scientific-program

TEHNIČKI GLASNIK / TECHNICAL JOURNAL – GODIŠTE / VOLUME 20 – BROJ / ISSUE 2

LIPANJ 2026 / JUNE 2026 – STRANICA / PAGES 173-342

Sveučilište Sjever

SVEUČILIŠTE SJEVER / UNIVERSITY NORTH – CROATIA – EUROPE

ISSN 1846-6168 (PRINT) / ISSN 1848-5588 (ONLINE)

TEHNICKIGLASNIK@UNIN.HR – HTTP://TEHNICKIGLASNIK.UNIN.HR